Improving crop modeling approaches for supporting farmers to cope with weather risks

DISSERTATION

zur Erlangung des akademischen Grades
Doctor rerum agriculturarum
(Dr. rer. agr.)

eingereicht an der
Lebenswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von
Herrn M.Sc. Christoph Gornott

Präsidentin der Humboldt-Universität zu Berlin
Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Lebenswissenschaftlichen Fakultät
Prof. Dr. Bernhard Grimm

Gutachter

1. Prof. Dr. Hermann Lotze-Campen

2. Prof. Dr. Reimund Paul Rötter

Tag der mündlichen Prüfung: 16. Februar 2018

Abstract

Due to changing climate and weather patterns in combination with limitations to extend global arable land area, the pressure on food production systems will increase. Additionally, these future food production systems must feed a rapidly growing world population, whose demand for food rises and becomes increasingly land-intensive. To cope with these challenges, it will be indispensable to increase and stabilize crop yields. This requires, however, a deeper understanding of the factors influencing crop yield variability and a quantification of their relevance under different soil and climate conditions. Besides crop field trials, crop modeling assessments are suitable methods to analyze such yield influencing factors. Therefore, however, these assessments need to be improved in order to appropriately cover the relevant factors influencing yield variability. This dissertation contributes to that research need as we[1] further develop and apply crop models to assess regional wheat and maize yield variability in Germany, Tanzania and on a global scale. For this, we analyze both statistical and process-based crop models in an intra- and inter-comparison and combine the advantages of both model types in a new modeling approach. We use both crop model types to decompose weather and non-weather-related crop yield variability and quantify the weather-related production risks for temperate and tropical production conditions. For achieving this, we apply five steps: (i) First, we develop a statistical crop modeling approach to decompose the influence of weather and agronomic management on winter wheat yields in Germany. (ii) Based on the first step, we expand the statistical methods and apply augmented models for winter wheat and silage maize on a disaggregated level. (iii) Then this model approach is used to investigate an out-of-sample cross validation to demonstrate the models' capability to project future yield changes under climate change. (iv) In a global statistical application, this models' capability of projecting yields is tested for short-term yield forecasts. (v) Finally, we combine statistical and process-based crop modeling to decompose weather-related maize yield losses from losses caused by non-weather factors for the case of Tanzania. Across these five steps, we find that the share of weather-related yield variability is higher in Germany than in Tanzania. Accordingly, crop yield variability in Tanzania is to a higher share attributable to agronomic management and socio-economic influences. For both countries, we find that the share of explained weather-related yield variability is higher on an aggregated level than on the regional level (i.e. districts, counties, or grid cells). This can be explained by heterogeneous management conditions across regions, which are averaged out by the spatial aggregation to national or subnational levels. Moreover, we demonstrate that our statistical models reproduce the observed yield variability well with a goodness of fit (R^2) mostly higher than 0.80 for Germany, Tanzania and globally. Furthermore, we are able to show that the statistical component of our approach can be used for short-term yield forecasts and to some extent also for climate change projections. Furthermore, the combined statistical-process-based approach can be used for assessing weather-related crop yield losses for insurance purposes. The application of crop models in yield forecast systems and insurance solutions could contribute to develop measures, which support improving food security on a global scale and notably in Sub-Saharan Africa.

[1]As the main text (chapter 2–6) of this dissertation contains other authors' contributions and for an easier readability, I will use "we" in the entire dissertation to refer either to me or to all contributing authors of the chapters 2–6.

Zusammenfassung

Sich ändernde Klima- und Wetterbedingungen in Verbindung mit einer begrenzt ausdehnbaren Ackerfläche werden den Druck auf Nahrungsmittelproduktionssysteme weiter erhöhen. Darüber hinaus müssen zukünftige Nahrungsmittelproduktionssysteme eine schnell wachsende Weltbevölkerung mit einer zunehmend landintensiven Nachfrage nach Nahrung ernähren. Um diesen Herausforderungen gerecht zu werden, ist eine Erhöhung und Stabilisierung der Ernteerträge unverzichtbar. Dies erfordert aber ein tieferes Verständnis der Einflussfaktoren, die auf die Ertragsvariabilität wirken, sowie die Quantifizierung ihrer Relevanz unter verschiedenen Boden- und Klimabedingungen. Neben Feldversuchen sind Ertragsmodellierungsansätze geeignete Methoden um die ertragsbildenden Einflussfaktoren zu untersuchen. Dafür müssen diese aber so weiterentwickelt werden, dass sie die relevanten Faktoren auf die Ertragsvariabilität besser abbilden. Diese Dissertation leistet einen Forschungsbeitrag zu Ertragsmodellen, die wir zur Abschätzung regionaler Weizen- und Maisertragsvariabilität in Deutschland, Tansania und auf globaler Ebene weiterentwickeln und anwenden. Dazu analysieren wir sowohl statistische als auch prozessbasierte Ertragsmodelle und kombinieren die Vorteile beider Modelltypen in einem neuen Modellierungsansatz. Somit verwenden wir beide Modelltypen, um wetter- und nicht-wetterbedingte Ernteertragsvariabilität getrennt zu analysieren und so die wetterbedingten Produktionsrisiken für gemäßigte und tropische Produktionsbedingungen zu quantifizieren. Um dies zu erreichen nutzen wir fünf Schritte: (i) Zunächst entwickeln wir einen statistischen Modellansatz, um den Einfluss von Wetter und agronomischem Management auf Winterweizenerträge in Deutschland zu separieren. (ii) Auf der Grundlage des ersten Modells erweitern wir die statistischen Methoden und wenden die erweiterten Modelle für Winterweizen und Silomais auf disaggregierter (Landkreis) Ebene an. (iii) Diesen erweiterten Modellansatz verwenden wir daraufhin zum Testen einer Kreuz-Validierung mit dem Ziel zukünftige Ertragsänderungen unter Klimawandel zu projizieren. (iv) Anschließend wird in einer globalen statistischen Anwendung die Kapazität dieser Modelle für kurzfristige Ertragsprognosen getestet. (v) Schließlich kombinieren wir für das Fallbeispiel Tansania statistische und prozessbasierte Ertragsmodelle, um wetterbedingte Ertragsverluste von nicht-wetterbedingten Ertragsverlusten zu separieren. Als Ergebnis der fünf Schritte lässt sich zusammenfassen, dass der Anteil der wetterbedingten Ertragsvariabilität in Deutschland höher ist als in Tansania. Dementsprechend sind die Ertragsschwankungen in Tansania eher auf das agronomische Management und sozioökonomische Einflüsse zurückzuführen. Für beide Länder stellen wir fest, dass der Anteil der wetterbedingte Ertragsvariabilität auf aggregierter Ebene höher ist als auf regionaler Ebene (Landkreise, Distrikte oder Gitterzellen). Dies lässt sich durch regional heterogene Managementbedingungen erklären, die durch die räumliche Aggregation zu nationalen oder sub-nationalen Einheiten herausgemittelt werden. Darüber hinaus zeigen wir, dass unsere statistischen Modelle die beobachtete Ertragsvariabilität mit Erklärungswerten (R^2) von meist über 0,80 für Deutschland, Tansania und weltweit reproduzieren. Wir können zeigen, dass der statistische Bestandteil unseres Ansatzes für kurzfristige Ertragsprognosen genutzt werden kann und teilweise auch für Klimawandelprojektionen nutzbar ist. Der kombinierte statistisch-prozessbasierte Ansatz zur Bewertung von wetterbedingten Ertragsverlusten kann für Versicherungszwecke genutzt werden. Die Anwendung der Ertragsmodelle in Ertragsprognosesystemen und Versicherungslösungen kann dazu beitragen Maßnahmen zu entwickeln, welche die globale Ernährungssicherheit vor allem in Afrika südlich der Sahara verbessern.

Acknowledgements

First of all, I would like to thank my supervisors, Hermann Lotze-Campen and Reimund P. Rötter, for supporting my work on this dissertation. In particular, I am especially grateful to Hermann Lotze-Campen and to my day-to-day supervisors Fred Hattermann and Frank Wechsung for their continuous and highly encouraging feedback. They shared their extensive knowledge and experience, which was very valuable while writing the dissertation and will further guide me through my future career. And as this cumulative dissertation is composed of jointly published journal articles, I am furthermore deeply grateful for the work of Hakon Albers, Fred Hattermann, Silke Hüttel, Bernhard Schauberger and Frank Wechsung. Without their contributions it would not have been possible to explore these research questions and publish the articles. Beyond the articles imbedded in this dissertation, I also highly appreciate the joint work on three further articles published or submitted with the corresponding authors Tobias Conradt, Marcos Lana, and Kristine Belesova.

I also owe gratitude to Peggy Gräfe for her positive, comprehensive and constant support.

The research for this dissertation is partially carried out within the Trans-SEC project (BMBF, BMZ), for which I would like to thank all partners. Notably, I am grateful for the discussions I had with Folkard Asch, Frieder Gräf, Jörn Germer, Ludger Herrmann, Christoph Müller, Nadja Reinhardt, Angela Schaffert, Stefan Sieber, and Siza Tumbo. Moreover, I highly appreciate receiving funding for the ongoing projects which I lead: YLIT (funded by Climate-KIC), CIMSU (funded by Climate-KIC), AgRATI (funded by Climate-KIC), H2020_Insurance (EU) [also led by Fred Hattermann, Tracy Irvine], GeoCare (BLE), Food security, migration and conflicts (Leibniz Association), CYE (Munich Re), Climate Risk Management in the Agricultural Sector of Peru (GIZ) and for the work, which has been done in these projects by Tobias Conradt, Sophia Rottmann, Bernhard Schauberger, and Michel Wortmann.

I also acknowledge the fruitful scientific discussions and conversation I had with my colleagues Valentin Aich, Timon Graf, Hagen Koch, Stefan Liersch, Anastasia Lobanova, Maria Martin, Christopher Reyer, Michael Roers, Sophia Rottmann, Judith Stagl, Kira Vinke, and Michel Wortmann. My special thanks go to my friends Andi, Felix, Hakon, Johanna, Lisa, Lukas, Olli, Paula, Simon, Srijna, and Theresa, as well as my family.

Finally and foremost, a heartfelt thanks goes to Sophia for supporting me throughout my dissertation. She encouraged me in times of doubt and shared with me many times of joy. I am truly grateful that she carefully proof-read my dissertation with a lot of patience and inspiration.

Improving crop modeling approaches for supporting farmers to cope with weather risks

Table of Contents

1 Introduction

1.1 Introduction and motivation

In the face of a changing climate, weather extremes will appear more frequently and severely and will make crop production more vulnerable (Challinor et al., 2007; Lobell et al., 2014, 2012; Semenov and Shewry, 2011). This will increase the pressure on temperate and tropical crop production systems and might have strong impacts on global food security (IPCC, 2014). In Low Income Countries, like Tanzania, variable crop yields directly affect food security, but also in High Income Countries, like Germany, changing crop yields influence global food supply (Foley et al., 2011; West et al., 2014; Wheeler and von Braun, 2013). To secure a stable and sufficient food supply, it is crucial to understand factors influencing crop yields across and within the different agro-ecological regions (Ewert et al., 2015; Liu et al., 2016). This understanding can support the development of coping strategies to forthcoming production risks, for example investments in risk reduction (e.g. irrigation techniques) and risk transfer solutions (e.g. crop insurances). By short-term yield forecasts or loss assessments, statistical crop models (Iizumi et al., 2013; Ray et al., 2015) and process-based crop models (Asseng et al., 2013; Bassu et al., 2014) can contribute to such risk reduction and transfer instruments. Moreover, these crop model assessments can underpin the data of crop field trials by investigating the yield influencing factors for different agro-ecological regions, thereby structuring the information, quantifying the relevance of the factors and upscaling the results to larger regions. This might have positive impacts on stabilizing and enhancing farmers' incomes and contribute to global food security (Lipper et al., 2014; Tilman et al., 2011; Wheeler and von Braun, 2013).

In many Sub-Saharan African (SSA) countries, actual crop yields remain significantly below the plant-physiological yield potential even though climate conditions (especially annual precipitation) would be sufficient to achieve this potential in many regions (van Ittersum et al., 2013). In these countries, the food security status is often classified as "serious" according to the Global Hunger Index (Wheeler and von Braun, 2013). Usually, this classification is associated with high numbers of people suffering from inadequate nutrition (hidden hunger) and chronic food insecurity. Due to climate change, it is likely that the group of affected people will further increase. Moreover, climate change induced food insecurity can amplify health problems (e.g. child mortality or water and vector borne diseases like malaria and cholera), violence (riots, armed conflicts) and unwanted migration (Kelley et al., 2015; Phalkey et al., 2015; Schleussner et al., 2016) and thus, impel the loop of food insecurity and poverty.

In this dissertation, we develop crop models and show the application for Germany, Tanzania and in a global approach. In the five main chapters, we demonstrate the application of crop models for decomposing and analyzing crop yield influences, climate change projections, short-term forecast, and loss determination for crop insurances. Each of the five main chapters is briefly introduced and motivated in by the following sub-chapters of this introduction.

1.2 Crop yield assessments

1.2.1 Decomposing yield impacts

Weather risks endanger agricultural production around the world. In particular in SSA, weather risks have strong impacts on crop production and food security, because often farmers do not have the capacity to adjust their agronomic management in case of weather extremes (Knox et al., 2012; Müller et al., 2011). In Europe, weather risks also affect agricultural production and – since Europe is an important food producer – global commodity prices. Thus, the European production influences global food availability and affordability. To gain knowledge of the effects of weather risks on food production, it is important to understand the influencing factors, which are responsible for crop yield variability. Crop models can be used to identify such factors. These crop models also allow decomposing different yield influences. This decomposing can be conducted by assigning individual yield influences such as precipitation or temperature (Fishman, 2016; Miao et al., 2016; Welch et al., 2010; You et al., 2009) or influence groups like weather, agronomic management, or economic impacts on yields (van Dijk et al., 2017). Understanding different sources of yield variability – due to the decomposing – can support farmers to adapt their agronomic management towards more resilient crop production. Moreover, the decomposing can be used to support risk transfer instruments like crop insurances and can support policy makers with information to counteract food crises or improve crisis management.

The decomposing of the crop yield influencing factors can be conducted with both process-based and statistical models. These two model types allow separating the weather-attributable impacts from the agronomic management-related yield impacts by adjusting the model in such a way that it only considers one of these two yield influencing groups. For that, process-based models require constant input data of one influencing group across the cropping seasons (Ewert et al., 2011; Folberth et al., 2016). In comparison, statistical models allow a simple decomposing by considering only one subset of regression parameters and its corresponding variables. This decomposing can be used for productivity assessments of single weather variables (You et al., 2009) and production risk assessments utilized by crop insurances or weather derivatives (Woodard and Garcia, 2008).

1.2.2 Influences on crop yield variability

Weather patterns determine and limit crop yields and influence its variability. The main weather influences on crop yields are atmospheric CO_2 content, solar radiation, temperature and crop water supply. While the – relatively constant – atmospheric CO_2 content and solar radiation rather determine the yield level, temperature and water supply are mainly responsible for crop yield variability. The key limitations are insufficient water supply – caused by the interaction of precipitation, evapotranspiration and soil properties – and non-optimal temperatures within the different development stages. Moreover, weather-related pests, weeds, and diseases further limit possible crop growth (Rötter and

Van de Geijn, 1999; Tittonell and Giller, 2013; van Ittersum et al., 2013). Statistical and process-based crop models are indispensable approaches to gain a deeper understanding of the factors influencing crop yield variability in different crop producing regions. Most statistical crop models include water availability and temperature as weather-related yield influencing factors (Butler and Huybers, 2012). Based on these two factors, several variables – like growing degree days, evapotranspiration or precipitation deficit – are used to account for the weather influence on crop yields. In addition, extremes of both weather factors are often considered in statistical models. These are droughts (Lobell et al., 2014), floods (Blanc, 2012; Rosenzweig et al., 2002), extreme heat (Lobell et al., 2013, 2011), and frost (Grassini et al., 2009).

Besides weather impacts, agronomic management and socio-economic factors also influence yield level and variability (van Dijk et al., 2017). While the agronomic management refers to directly applied measures like fertilizer application (van der Velde et al., 2014), irrigation (You et al., 2011) and other production factors (You et al., 2009); the socio-economic factors include the acreage (Iizumi and Ramankutty, 2015), prices (Miao et al., 2016) or subsidies (Sánchez, 2010). The socio-economic factors influence the use of agronomic management measures and thus, indirectly impact crop yields. In the European Union (EU), crop yields are mostly achieved with sufficient input supply, while it is mostly insufficient in SSA (Tittonell and Giller, 2013; Vitousek et al., 2009). The insufficient and unbalanced application of (nitrogen and phosphorus) fertilizer is often the reason for low crop yields (van der Velde et al., 2014). Moreover, the application of other inputs like plant protection measures also highly differs in SSA (Christiaensen, 2017) and thus, affects crop yields. Besides the direct influences on crop yields due to agronomic management, indirect impacts influence the farmers' behavior. For instance, input subsidies – as they are largely disbursed in SSA – have an impact on the economic return of input usage and thus, may change farmers' input use, which in turn has a direct impact on crop yields. The Common Agricultural Policy (CAP) of the EU has decoupled the subsidies from the production (area-based direct payments). Its aim is to prevent trade-distorting effects caused by the subsides (Gohin, 2006; WTO, 2017). In addition, other socio-economic factors impacting farmers' behavior may influence crop yields. These factors are, for instance, input and commodity prices in the EU and factors such as market access, land tenure security or access to extension services in SSA.

1.2.3 Yield impacts of climate change

Due to rising global temperatures and changing precipitation patterns, longstanding agronomic practices have to be adjusted in order to retain current crop production levels in different world regions (Foley et al., 2011; West et al., 2014). Climate change simulations in combination with crop models can be used to project possible future crop yields. Such yield projections can support farmers' decision making or the scope of plant breeding in regard to long-term climate change adaptation strategies.

Globally, the temperatures will further increase as projected by all climate simulations models (IPCC, 2014). Their projections also show that annual precipitation will – depending on the world region – either increase or decrease (see Fig. 1). Due to rising temperatures, the absolute water holding capacity of the atmosphere and hence atmospheric water demand (also called potential evapotranspiration) and the within-season precipitation variability will increase. The latter will occur because of an augmented probability for seldom, but heavier precipitation events. In most of the current cropping regions and notably in the tropical regions, the projected temperature will exceed the temperature, which is optimal for crop growth, by the mid of the 21st century (IPCC, 2014). This will add to the possibly negative crop yield impacts of higher potential evapotranspiration and within-season precipitation variability. Apart from these direct climate-related yield impacts, indirect impacts of changing climate conditions may also influence crop yields. These are, for instance, degraded soils through rainfall-induced erosion or increasing pressure of pests and diseases through changing climate patterns (Rosenzweig et al., 2001; Sileshi et al., 2010).

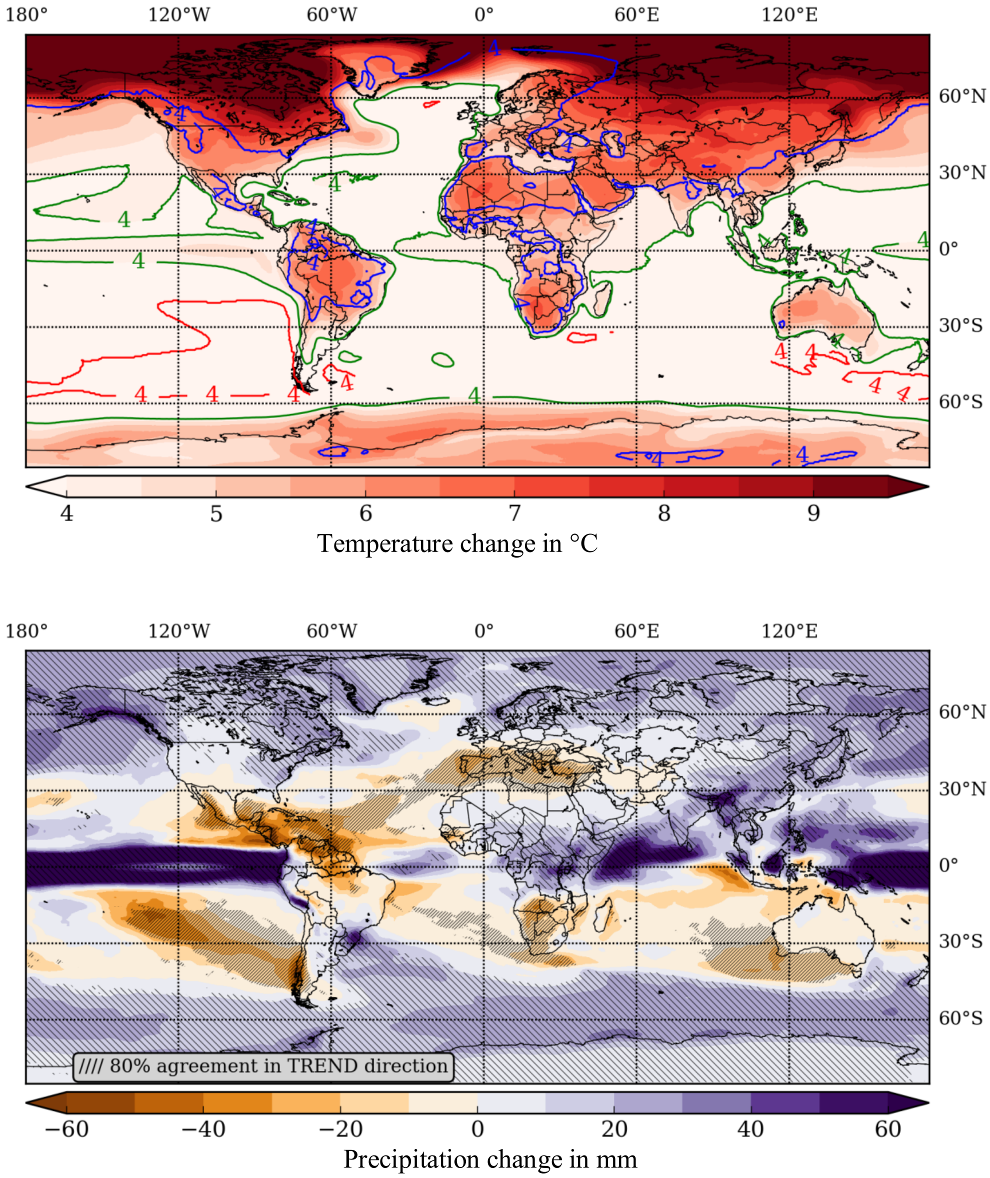


Fig. 1: Projected trends in temperature (top) and precipitation (bottom) under the high-end Representative Concentration Pathway (RCP) with an increase of radiation by 8.5 W/m² by 2100 (RCP 8.5). Both plots show the mean trend of 18 model runs of the Coupled Model Intercomparison Project Phase 5 - Global Circulation Model (CMIP5 GCM) ensembles. The temperature increase is shown in °C and the precipitation change in mm (per year) between 2006 and 2100. The shaded areas show where at least 80% of the climate simulations agree on the trend. The RCP 8.5 represents the high-end, but most-likely scenario under the current emission pathway. The analysis (including this figure) was conducted by Peter Hoffmann (PIK-Potsdam).

Crop models can be used to project crop yield impacts of changing climate conditions. This can support farmers to stabilize (and enhance) crop yields and cope with uncertain climatic conditions in the future. Process-based crop models are widely used to project these impacts of climate change on future crop yields (Folberth et al., 2012; Rosenzweig et al., 2014). These models project climatic yield impacts beyond the observed range of yield and weather variability due to its bio-physical organization (Ewert et al., 2015). However, these process-based models face the problem that they have to use frequently biased climate simulation data (Müller et al., 2016). Because of the great efforts needed to correct such biases (Hawkins et al., 2013; Lobell, 2013), approaches are needed, which do not require or are able to overcome the complex procedure of bias correction. Moreover, process-based crop models should also allow for projecting crop yield responses caused by extreme temperatures, droughts and extreme precipitation, which were not observed in the past.

1.2.4 Seasonal forecasts

Crop yield forecasts, which identify yield losses within the current growing season, are of high interest to support farmers' agronomic and risk management. If the information about possible crop yield losses or failures is available before they occur, it will help farmers to adjust their agronomic management by implementing counteracting measures (e.g. irrigation, adjust/control harvesting dates by agronomic management measures). Moreover, it would corroborate policy makers' decisions on providing financial reliefs to support affected farmers prior to or immediately after a yield loss (Qian et al., 2009; Stone and Meinke, 2005). Despite the high interest for such forecasts by farmers, retailers, insurance companies, and other stakeholders along the food value chain, the accuracy and spatial coverage of the available forecast systems differ highly. So far, there is no worldwide forecast system with a standardized approach, which meets the required accuracy. To assess forthcoming production shortages, crop model approaches can be linked with weather forecast data to provide seasonal yield forecasts as done for the EU in the Monitoring Agricultural ResourceS (MARS) project (MARS, 2017), for Canada in the Integrated Canadian Crop Yield Forecaster (ICCYF) (Chipanshi et al., 2015), or for the US and other world regions by the National Agricultural Statistics Service (NASS) of the United States Department of Agriculture (USDA, 2017). If such forecasts were available at a global scale, it would be possible to link the forecasts with commodity price models and thus gain insights on upcoming price changes (Schewe et al., 2017). Furthermore, such forecasts can help to alleviate weather risks in crop production and to choose adequate strategies of risk-adjusted input intensity (Berg et al., 2009; Stone and Meinke, 2005). Since crop production and commodity prices affect food security, these forecasts can contribute to enhance local food security (Gilbert et al., 2017; Wheeler and von Braun, 2013).

1.2.5 Loss assessments for insurance schemes

Micro-insurance solutions are often presented as important tools to enhance resilience to climate change and altered weather perils in SSA (IPCC, 2014; Surminski et al., 2016). Such insurance solutions can help transferring the risks from smallholder farmers to other sectors like the finance sector. The latter is more apt to deal with these systemic risks (Conradt et al., 2015), and thus the finance sector can contribute to stabilize smallholder farmers' incomes. Moreover, such insurances do not only indemnify the economic value of yield losses, but can also create other co-benefits for smallholder farmers (see Fig. 2). Possible co-benefits are for instance enhanced food security, indemnified livelihoods, positive impacts on smallholder farmers' health and lives in general (Meze-Hausken et al., 2009). In case of weather-related yield losses, insurance claims can help to purchase food (co-benefit enhanced food security) and prevent that farmers loose or have to sell their agricultural inventory (co-benefit livelihoods' indemnification) in years of extreme yield losses. Since smallholder farmers have to advance money to purchase seeds and other agricultural inputs for the next growing season, there is often less or no money to invest in agricultural inputs and production techniques for enhancing crop yields after a crop failure. Furthermore, increased income stability, due to the insurance claims, will give farmers higher creditworthiness, because banks then consider reduced loan repayment risk. And thus, access to micro-credits allows investing in agricultural inputs and production techniques, whose purchase is too risky without the insurance. This can unlock a positive agricultural adaptation loop achieving higher overall crop yields and food security, livelihood indemnification and resilience and might further enhance farmers' ability to adapt to changing climate conditions (Cole et al., 2013; Meze-Hausken et al., 2009). This can increase the resilience of crop production systems.

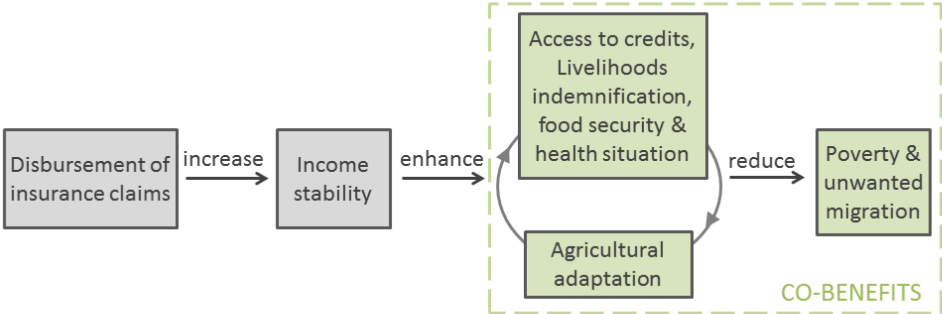


Fig. 2: Potential impact chain of insurance solutions and corresponding co-benefits for smallholder farmers.

In SSA, farmers largely lack sufficient financial capacity to adjust their agronomic management when extreme, unexpected weather conditions occur. In the face of increasing weather extremes due to climate change in combination with low stockpiles, smallholder farmers are very vulnerable towards severe yield losses. Crop insurances can help reducing this risk in crop production. However, a widespread implementation of insurance schemes is hindered by inaccurate and unavailable yield loss determination as well as by high costs for these determinations. While index-based insurance solutions often face the problem of a low accuracy of the loss determination, indemnity-based insurances require claim adjusters to determine insurance payouts. Because of small farm sizes and underdeveloped road systems in remote SSA regions, loss determination with the help of claim adjusters is very expensive. It would rise transaction costs of potential insurance schemes and thus, make it unaffordable for smallholder farmers in SSA. If weather-related crop yield losses were accurately assessed at affordable costs, it would be possible to implement index insurances in regions with only regional knowledge about yields and respective losses. As such assessments are the aim of crop modeling since the 1960s (Oury, 1965; Ritchie, 1972), statistical and process-based crop models can contribute to determine yield losses for insurances purposes (Finger, 2013; Linnerooth-Bayer et al., 2011). An increased loss assessment accuracy will help to build trust among farmer and insurance, rise farmers' willingness to pay and secure a long-term and sustainable implementation of insurance schemes (Conradt et al., 2015; Hill et al., 2013).

1.3 Methodical approaches

Commonly crop models are assigned to two main approaches: statistical and process-based. While statistical crop models generally use regression approaches to reproduce observed yield data, process-based models use physically-based algorithms to calculate the impacts on crop yields without using observed yield data for the calculations. To some degree, however, the differentiation between both model types is indistinct and the approaches overlap in some points. For instance, statistical models often use pre-processed weather variables (e.g. potential evapotranspiration), while process-based models often also contain information of empirical approaches, for example linear relationships between environmental variables. Nevertheless, these two model types calculate crop yields with different approaches and thus, have different advantages and disadvantages.

1.3.1 Statistical models

Statistical crop models estimate the impact of yield influencing (exogenous) variables – within a predefined functional form – on the endogenous variable crop yield. The exogenous variables are either only weather factors (see for example the models developed by Blanc, 2012; Lesk et al., 2016; Ray et al., 2015) or a combination of weather and non-weather factors. The latter comprise factors of agronomic management and socio-economy. Such models have been developed by e.g. Ward et al. (2014) and You et al. (2009). In these models, weather data is aggregated over an entire period or sub-periods of a growing season to capture the direct weather influences on crop yields, but also collinear (indirect) impacts of weather on crop yields. Such indirect effects are for instance the occurrence of pests, weeds, and diseases. Since this information is included in the observed yield data, statistical models implicitly control for these indirect influencing factors. Statistical models can also consider the impacts of agronomic management and socio-economy. These indirect socio-economic impacts – which influence agronomic management – can be considered as a proxy for unknown management conditions. This is an important advantage of statistical models in particular in regions with limited data availability. Finally, time-invariant yield impacts like soil quality can be captured with statistical models in the constant term of a linear regression (intercept) or due to the variable transformation (fixed effects or first differences).

1.3.2 Process-based crop models

Apart from statistical models, process-based models are an indispensable tool for analyzing yield impacts of changing weather and agronomic conditions. These models compute the impact of weather, soil, and agronomic management conditions on crop yields with by interacting sub-processes resolved mostly in daily time steps. These sub-processes are for instance biomass growth, photosynthesis, transpiration, nutrient uptake, plant development, soil dynamics, and other plant-physiological relevant functions. Process-based models are applicable to a large range of environmental conditions without changing the parameters (Asseng et al., 2013; Bassu et al., 2014), which are mostly observed in crop field trials. In contrast, the regression parameters of statistical crop models vary stronger across regions and might be only valid within the range of the observed environmental conditions. Process-based models ideally can project yield impacts beyond the observed range of e.g. extreme temperatures, droughts, dry spells, growing season shifts (Thornton et al., 2011, 2009) and nutrient shortages (van der Velde et al., 2014). However, these models are mostly restricted to processes directly observable at plot or farm level and do not integrate impacts of socio-economic conditions.

1.3.3 Combining statistical and process-based crop models

Largely, either process-based or statistical models are applied for crop yield assessments. In some cases both model types are compared (see e.g. Estes et al., 2013; Liu et al., 2016; Lobell et al., 2005; Lobell and Burke, 2010). Both model types have advantages and disadvantages, however, so far there is no approach which connects both model types. Since the advantages of one model type are often the weaknesses of the other model type, a combined approach – which makes use of both models' advantages – offers an opportunity to increase the robustness of yield assessments and projections (Rötter et al., 2011). One main advantage of process-based models is that they can integrate the impacts of different agronomic management conditions and can project in yield levels, which are not observed in the past (Asseng et al., 2013). In contrast, statistical crop models have the advantage to allow decomposing yield variability in weather-related and socio-economic yield variability (You et al., 2009).

1.4 Case studies

Around the world, weather and climate conditions, quantity and quality of available land resources as well as agronomic management measures determine crop production and its annual and spatial variability. To further investigate and understand the different impacts influencing crop yields, we select two countries with different climate and input levels to carry out case studies. We select Germany as one country with temperate climate and high-input agricultural systems and Tanzania with tropical climate and low-input agricultural systems (Fig 3). We then compare the two regions with respect to their yield influencing factors and weather-dependency of crop yield variability. Furthermore we conduct a global analysis.

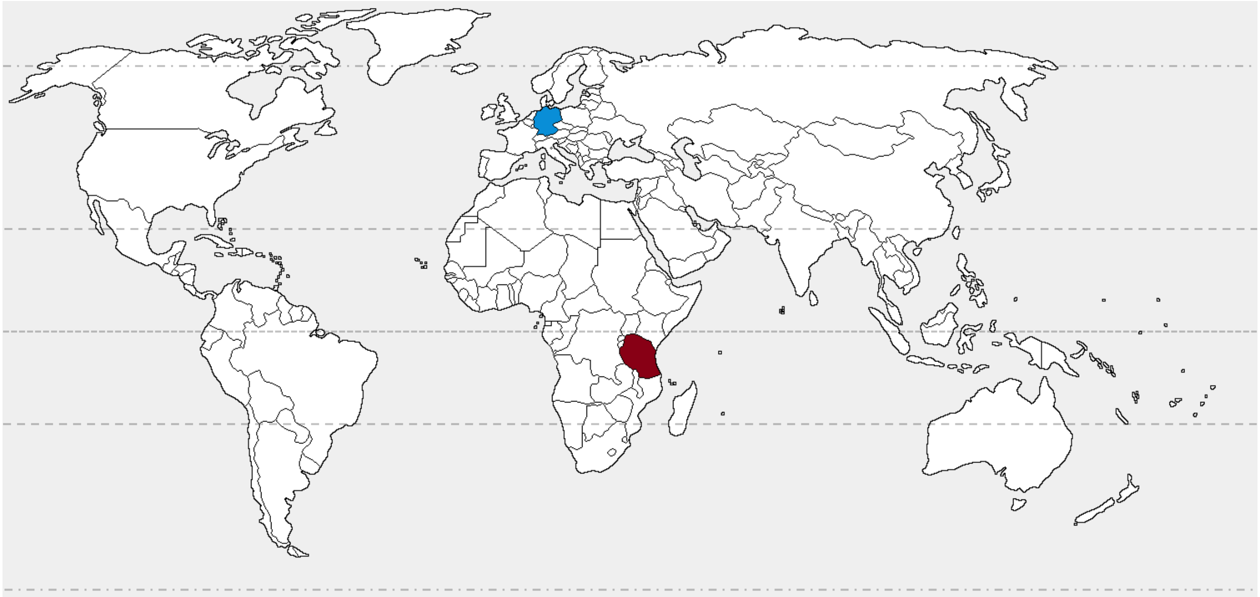


Fig. 3: World map with the location of the two case studies Germany (blue) and Tanzania (red). The equator, the two Tropical Circles and Polar Circles are shown as dashed line.

1.4.1 Climate conditions

Germany is characterized by a temperate climate with a more maritime climate in the Northwest (in particular at the coast) and a continental climate in the East. The average summer temperature (JJA) from 1951 to 2010 was 17.8 °C and in winter (DJF) 2.8 °C. Average annual precipitation mostly ranges between 500 and 900 mm. At the Bavarian Alps, annual precipitation usually averages above 2,000 mm. Between 1951 and 2010, the lowermost annual precipitation was 215 mm measured in Saxony–Anhalt and the highest with 3,503 mm measured in the Alps (own calculations, data: DWD weather station data, 2012). Precipitation is evenly distributed over the whole year but has its maximum monthly rainfall in summer and the minimum monthly rainfall in winter in most areas of Germany. Tanzania has a tropical and sub-tropical climate with a dry and rainy season. In the south–western lowlands average annual precipitation ranges from 700 to 2,000 mm and in the northern semi-arid highlands from 400 to 700 mm. Monthly average temperatures range between 18 and 28°C throughout the year (own calculations, data: WFDEI of Weedon et al., 2014). Thus, the cropping conditions are characterized by high spatial and temporal weather-induced heterogeneity (Ramirez-Villegas and Challinor, 2012; Rowhani et al., 2011). These diverse climate conditions qualify Germany as a good study region for cropping conditions in Europe – where annual precipitation ranges between 300–2,000 mm – and Tanzania for SSA – where the annual precipitation range is 200–2,500 mm.

1.4.2 Land use types and farming systems

Germany has a total land surface area of 35.7 million ha. Agriculture covers 48% of this area; this includes 34% (11.8 million ha) used as arable land (average 2001–2015). Since the 1960s, the arable land is constant, but the number of farms decreased (to some 275’000 in 2016) and farms tend to get more specialized (German Federal Statistical Office, 2017) . In contrast, Tanzania has a total area of 94.5 million ha with 44% of the total land area are used for agriculture and 15% of this total land (14.2

million ha) are arable land (World Bank, 2017). Between 2003 and 2012, the arable land area increased by about 50% (World Bank, 2017) and still there is substantial potential for enlargement. Fischer and Shah (2010: 10) estimate that the arable land could further increase by up to 50%. In Tanzania, agricultural land is cultivated by 10 million smallholder farms (FAO, World Bank, 2016). Under the Tanzanian Village Land Act (from 1999), the village (and not the smallholder farmers themselves) is the primary land-holding unit and responsible for land administration, land-disputes settlement as well as the recognition of customary land tenure and transferring it to formally-granted land rights (Knight, 2010). However, only a few Tanzanians are informed about these tenure rights (World Resource Institute, 2010).

In Germany, the most planted crops are winter wheat (*Triticum aestivum* L.) and silage maize (*Zea mays* L.). Between 1991 and 2010, the silage maize acreage increased by 50% from 1.3 to 1.8 million ha and the grain maize acreage by 100% from 0.23 to 0.46 million ha. The wheat acreage (which contains 98% winter wheat) increased in the same period by 35% from 2.4 to 3.3 million ha (the latter uses 28% of total arable land). Other important crops are winter barley with an acreage of 1.4 million ha, canola with 1.1 million ha as well as spring barley and ray with 0.7 million ha each (Statistical Offices of the Federation and the Länder, 2016) . In Tanzania, the most planted crop is maize (*Zea mays* L.) covering on average 3.0 million ha between 2000 and 2015. This is followed by cassava (1.0 million ha), beans (0.9 million ha), rice (0.8 million ha), and sorghum (0.7 million ha) (FAO Country STAT, 2017 corroborated by FAO Stat, 2013). In particular maize and legumes (mostly pigeon pea, groundnut or cowpea) are widely cultivated in intercropping systems (Snapp et al., 2014).

German agriculture is an input-intensive production system (Grassini et al., 2013), while Tanzanian agriculture is characterized by smallholder farming with insufficient access to agricultural inputs like fertilizer. According to Vitousek et al. (2009) and Tittonell and Giller (2013), the fertilizer usage is very low in SSA and rather poor than sufficient. In particular in eastern Africa, smallholder farmers apply insufficient and unbalanced (between the nutrients) amounts of fertilizer. Thus, nutrient withdrawal by harvesting exceeds nutrient replenishment through fertilizers (van der Velde et al., 2014). If this nutrient gap is closed, this will result in a large maize yield increase (Mueller et al., 2012). In Tanzania, between 2003 and 2010 the average actual maize yield is only 1.3 t ha^{-1} (MAFSC, 2010). Nevertheless, field trials show that Tanzania has a high potential to increase maize yields (van der Velde et al., 2014). Within the same period, Germany achieved on national arithmetic average grain maize yields of 8.9 t ha^{-1} (FAO Stat, 2013) , winter wheat yields of 7.1 t ha^{-1} and silage maize yields of 44.6 t ha^{-1} (Statistical Offices of the Federation and the Länder, 2016).

1.5 Structure of the work

After the introduction (1), this work is structured in five chapters and each chapter is one of the five published and peer-reviewed articles, which were written in collaboration with different (co-)authors during my PhD program.

(2) *Albers, Gornott, Hüttel, 2017: How do inputs and weather drive wheat yield volatility? The example of Germany, Food Policy (70) 50-61. doi: 10.1016/j.foodpol.2017.05.001*

In the first article, we investigate the applicability of a statistical crop model for the German wheat production on an aggregated level by employing a large set of weather and agronomic management variables. We investigate different exogenous variables and functional forms to describe the relationship between crop yield and yield impacting factors. Finally, we decompose weather and non-weather (agronomic management and economic factors) yield impacts to derive the most important yield impacting factors for wheat production in Germany.

(3) *Gornott & Wechsung, 2015: Level normalized modeling approach of yield volatility for winter wheat and silage maize on different scales within Germany, Jornal für Kulturpflanzen (67) 6, 205–223. doi: 10.5073/JFK.2015.06.01 [main text in German]*

In the second article, we use a similar statistical crop model approach, but on lower spatial scale (i.e. counties). We used the model to analyze winter wheat and silage maize yields. To make the results comparable with the first article and to increase the model robustness, we aggregate the county yields to federal states in a post processing. Here, we apply and discuss the setup of two further statistical regression methods to explain yield variability on a regional and aggregated level in Germany. In this article, we test different aggregation levels and discuss the impacts of water availability on yield variability. In this article, the different types of used regression models and variable selection are examined and discussed.

(4) *Gornott & Wechsung, 2016: Statistical regression models for assessing climate impacts on crop yields: A validation study for winter wheat and silage maize in Germany, Agricultural and Forest Meteorology (217) 89–100. doi: 10.1016/j.agrformet.2015.10.005*

The third article investigates and discusses the performance of the different statistical models considered in the second article. We do this with an out-of-sample cross validation and furthermore test the assessments' robustness and their applicability for climate projections on regional and aggregated scales. Moreover, the impacts of weather and non-weather influences on winter wheat and silage maize yields are shown on regional and aggregated levels.

(5) *Schauberger, Gornott, Wechsung, 2017: Global evaluation of a semiempirical model for yield anomalies and application to within-season yield forecasting, Global Change Biology, 1-15. doi: 10.1111/gcb.13738*

In the fourth article, we extend the existing statistical models and apply it to grain maize, soybean, spring and winter wheat at a global scale. We demonstrate their usage for climate change projections, which is discussed in the second article. Finally, we investigate the model performance to forecast crop yields one or two moth prior to the scheduled harvest time.

(6) *Gornott, Hattermann, Wechsung: Covering smallholder farmers' weather perils – a crop model based insurance approach for Tanzania, In review.*

In the last article, we combine statistical and process-based crop yield models and show the applicability of this combined approach for index-insurance solutions in Tanzania. In this article, we investigate several combinations between these two models types and discuss the application for loss determination and its uncertainty. Moreover, we calculate the premiums' costs for the Tanzanian maize production and setup a framework for a potential insurance implementation scheme.

Finally, this work closes with a general discussion about all five articles (7), conclusion (8), and an outlook for further research and application possibilities for such crop yield models (9). The structure of the general discussion is aligned to the five main chapters.

Summary articles:

(1) Albers, H., Gornott, C., Hüttel, S. 2017: How do inputs and weather drive wheat yield volatility? The example of Germany, Food Policy (70) 50–61. doi: 10.1016/j.foodpol.2017.05.001

(2) Gornott, C., Wechsung, F. 2015: Level normalized modeling approach of yield volatility for winter wheat and silage maize on different scales within Germany, Jornal für Kulturpflanzen (67) 6, 205–223. doi: 10.5073/JFK.2015.06.01

(3) Gornott, C., Wechsung, F. 2016: Statistical regression models for assessing climate impacts on crop yields: A validation study for winter wheat and silage maize in Germany, Agricultural and Forest Meteorology (217) 89–100. doi: 10.1016/j.agrformet.2015.10.005

(4) Schauberger, B., Gornott, C., Wechsung, F. 2017: Global evaluation of a semiempirical model for yield anomalies and application to within-season yield forecasting, Global Change Biology, 1–15. doi: 10.1111/gcb.13738

(5) Gornott, C., Hattermann, F., Wechsung, F.: Covering smallholder farmers' weather perils – a crop model based insurance approach for Tanzania, In review.

2 How do inputs and weather drive wheat yield volatility? The example of Germany

Hakon Albers[1*], Christoph Gornott[2], Silke Hüttel[3]

[1] Martin Luther University Halle-Wittenberg

[2] Potsdam Institute for Climate Impact Research (PIK)

[3] University of Rostock, Agricultural Economics

[*] Corresponding author

2.1 Abstract

Increases in cereals production risk are commonly related to increases in weather risk. We analyze weather-induced changes in wheat yield volatility as a systemic weather risk in Germany. We disentangle, however, the relative impacts of inputs and weather on regional yield volatility. For this purpose we augment a production function with phenologically aggregated weather variables. Increasing volatility can be traced back to weather changes only in some regions. On average, inputs explain 49% of the total actual wheat yield volatility, while weather explains 43%. Models with only weather variables deliver biased but reasonable approximations for climate impact research.

Keywords: Yield Wheat Variability Risk Weather Common Agricultural Policy

2.2 Introduction

Climate change and its consequences for agricultural production have been open to environmental, social and economic debate for years. This is not surprising since weather conditions consider- ably determine crop yield levels and their variability, which are of interest for food security reasons at the macro-level (Brown et al., 2015; Wheeler and von Braun, 2013). Yields are also interesting at the micro-level, where a low level of yearly crop yield variability reduces income risks and contributes to farm income stability, which in turn could be relevant at the macro-level in that it war- rants resilient food production. Hence, it is vital to better understand what determines yield variability in the most important crop producing regions. This may also help farmers adapt their agronomic strategy towards better-known risks, and help policy makers to prevent food crises or improve crisis management.

Undisputedly, long-term climatic changes alter cropping conditions (Siebert and Ewert, 2012) and might already have affected crop yield variability, which is identified as a key production risk of the most economically important cereals (IPCC, 2014, p. 71). Extreme weather events like the European heatwave in 2003 were discussed as either indicating an increase in temperature variability or resulting from a shift of the temperature distribution (Luterbacher et al., 2004; Perkins, 2015; Schär et al., 2004). Consensus exists that in the future, extreme weather events are expected to occur with greater frequency and severity in both temperate and tropical regions (IPCC, 2014, p. 69–73). This will likely make crop production more vulnerable, with potentially considerable impacts on farm incomes and food security, particularly in less developed regions.

Farmers can control inputs like fertilizer for a given natural production environment like soil quality but cannot control the weather, nor can they affect developments in markets, agricultural, or environmental policy. Weather[2] is exogenous to farmers and directly affects crop yields. Additionally, indirect effects entailing input adjustments exist. For instance, weed growth, pests and diseases vary depending on weather conditions and farmers usually adjust their inputs accordingly during the production period. Weather can be interpreted as the major driver of production risk in crop production, though the question remains, how much overall production risk can actually be traced back to changing weather conditions?

In this study we consider wheat – one of the most important cash crops worldwide – where considerable upward trends in both yield levels and variability have been observed. While in 1995/96, on average, about 2.5 metric tons per hectare (tons ha^{-1}) were harvested worldwide, in 2012/13 this increased to about 3.2 tons ha^{-1} (FAOstat, 2015). Our investigation concentrates on Germany, which produces 17% of the European Union's (EU) wheat output. In the period 1995/96 to 2012/13, German wheat yields increased from 7.1 to 7.7 tons ha^{-1}. Although a long period of relative yield stability existed in the 20th century (Calderini and Slafer, 1998), both absolute and relative wheat yield variability have increased in Germany since the 1990s (Krause, 2008; Osborne and Wheeler, 2013). Particularly concerning is the upward trend in relative yield variability, that is, an increased proportion of yield at risk relative to the expected mean.

Against this background, the research questions guiding our analysis are as follows: How to explain increasing relative yield variability? Particularly, can one really conjecture that production risk measured as relative yield variability has increased only through changes in weather conditions, as the climate change discussion implies?

Several other reasons for this increase exist. First, farmers might adjust input levels because of changing input and output prices (Miao et al., 2016), while Finger (2010) discussed the importance of agricultural policy for yield analyses. Farmers in the EU have been exposed to rather radical changes in

[2] We use the term "weather" to be consistent with the majority of papers we reviewed. The literature applies different definitions. Dell et al. (2014) refer to inter-annual weather variations as long as the aggregation period is less than one year. Another strand of literature favors using a year-to-year or inter-annual variation of "climate" (e.g., Ray et al., 2015). Gornott and Wechsung (2015, 2016) use also the term "climate". In these chapters, we replaced the term "climate" by "weather" to make it congruent with the other chapters of this dissertation.

the Common Agricultural Policy (CAP) since 1992. Several reforms elevated the relative competitiveness of wheat, for instance, by removing price support, subsidies and compulsory set-asides (e.g., Gohin, 2006). Addition- ally, renewable energy policies have been proven to favor maize for silage (in Germany, increases of about 21% in the years 1990– 2009 were reported, Statistisches Bundesamt, 2015). This might also have contributed to changes in the relative competitiveness of wheat production, which has consequences for input intensity and thus crop yield levels (Banse et al., 2008; Schulze Steinmann and Holm-Müller, 2010). Overall, these policy changes may have provided incentives for farmers to use lower quality (marginal) land for wheat production, likely with negative effects on average yield levels and increased variability. Crops planted on marginal soils with low water-holding capacity might be more sensitive to extreme temperature and precipitation changes compared to more favorable soils (Perkins, 2015, p. 248–249). Moreover, yield can be interpreted as land productivity and may have increased due to scale and specialization effects (e.g., Yang et al., 1992; Kaufmann and Snell, 1997). Ongoing consolidation processes in the EU's agricultural sector (i.e., increased farm sizes) might enhance average yields per hectare despite the growing trend of planting marginal land with wheat.

While numerous studies consider how weather interacts with crop yield levels and their variance based on regression models (e.g., Chen et al., 2004), the relation between weather and relative yield variability of non-experimental yields has been analyzed by few researchers, for instance, Lobell (2007) or Ray et al. (2015). These authors, however, do not acknowledge any input adjustments that influence yield stability. To the best of our knowledge, thus far, the sources of yield volatility have not been disentangled into the major drivers of weather and inputs. Within this study we aim to close this gap and illustrate this idea using a case study for wheat yields in Germany.

While Iglesias and Quiroga (2007) assess the impact of weather variables on crop yields using time series regressions, we apply a panel data approach. We exploit the advantages of the panel structure to quantify whether and how weather- and input-induced risk has changed overall or only in some parts of Germany over time. Within our approach, we augment the contribution from Osborne and Wheeler (2013) and show that both inputs and weather matter for explaining yields and their relative variability. Our research contributes to the discussion of whether inputs need to be modeled when assessing climate change impacts on cereal yields. Further, understanding how weather drives observed relative yield variability today might be helpful for future adaptation challenges. Our empirical analysis involves two major steps. First, we develop an empirical model of relative yield variability consistent with a production function approach. We consider major inputs, test for suitable functional forms and enhance this production function by a rich set of weather variables addressing phenological development. Second, we decompose the fitted values of this regression model to disentangle weather-induced compared to input- or policy-induced relative yield variability referring to the approach by You et al. (2009). To improve our understanding of whether to control for input adjustments while relating weather and yields, we present an alternative model that leaves out major inputs. Hypothesizing that the latter may suffer from omitted variables bias, our results show no considerable qualitative differences, though they do exhibit quantitative differences.

In what follows, we first unfold the conceptual framework and present related literature. After introducing the data, the presented framework leads us to our empirical strategy for disentangling crop yield volatility drivers. Following that, we report and discuss our results, and finally conclude.

2.3 Conceptual framework and related literature

Numerous studies deal with the impact of weather on yield levels by using either process-based crop simulation models (Müller and Robertson, 2014) or regression techniques.[3] The latter approach finds its roots in Oury (1965) and has two major strands. First, many studies exist that simply relate yield and weather within a regression model (e.g., Butler and Huybers, 2015; we refer here to the literature overview Tab. S3–S5 in the supplementary appendix [SA]). In the second strand, weather impacts are analyzed within a production function framework including inputs. These models treat weather exogenously; however, a need to adjust inputs to changing weather might exist. For instance, the precipitation level will affect fertilizer intensity. Temperature instead affects length of the growing season and as such contributes to yield levels but rarely induces short-run adjustments to the input mix. While the first group of models takes this tacitly as a motive for leaving out inputs, the second strand of literature can also be criticized. While accounting for adjustments in the input mix in the short-run, production functions often fail to capture long-term adaptations to changes in climate such as altering crop rotation or alternative land-uses (e.g., Mendelsohn et al., 1994 or Deschênes and Greenstone, 2007).

When hypothesizing yield to be a function of inputs and weather, neglecting one group in the estimation of the impact of the other could result in biased parameter estimates as discussed by Kaufmann and Snell (1997), Reidsma et al. (2007, p. 417) or more recently by Miao et al. (2016, p. 201). In light of this debate, rather surprisingly only few recent studies include inputs or acknowledge other economic variables while analyzing weather impacts on yields (e.g., among others Schlenker and Lobell, 2010; Lobell et al., 2011; Blanc, 2012 or Ward et al., 2014). In this context, scale effects with regard to land have also been shown to influence yield levels (e.g., Chen et al., 2004). Hence, we rely on a production function approach including major inputs.

[3] Literature reviews can be found in Dell et al. (2014), Schlenker and Roberts (2009), Tannura et al. (2008) and Ward et al. (2014).

Disentangling the impacts that weather and inputs have on crop yield levels and their volatility, however, remains a challenge (You et al., 2009, p. 1013). Technically, a variety of approaches exist that quantify weather effects in a production function framework. We identify three crucial choices: the *selection* of weather variables, *aggregation* levels of weather data and the *functional form* describing the input-output and weather-yield relationships (further discussion of these choices in the SA).

Using aggregated data allows us to isolate the systemic component of weather risk at the federal state level simply because idiosyncratic shocks evident at the farm level are "averaged out" at higher aggregation levels (Marra and Schurle, 1994, p. 69; Woodard and Garcia, 2008). On the other hand, using aggregated data includes the disadvantage of a loss of information. Climate impact research typically works at lower levels though it focuses on identifying location-specific impacts under climate change.[4] In addition, we acknowledge that statistically more advanced and flexible ways to model systemic risk in yields or weather exist, for instance copulas (e.g., Gaupp et al., 2016; Xu et al., 2010). Our approach, however, targets at disentangling how, in addition to weather, inputs, policy, and macroeconomic shocks specifically drive wheat yields. As such, we connect insights from risk and productivity analysis, agronomic, and climate impact research.

2.4 Data

In what follows we describe the variables for the production function, yields and inputs, followed by weather and phenological stages (all details in the SA).

2.4.1 Production function for wheat

We analyze 12 German federal states[5] for the years 1995–2009. To specify the production function at the regional levels, we use accounting data from the European Farm Accountancy Data Network (FADN) provided by the European Commission (European Commission, 2015). These data contain representative farms from a stratified, rotating sample (Barkaszi et al., 2009). We refer to published results aggregated at the federal state level and select specialized crop farms referring to the EUs classification (i.e., specialist field crops according to the TF-8 grouping).[6] Our sample represents, on average, 4344 farms per state.

We specify the production function with one output (wheat yield) and eight inputs: capital, labor, wheat acreage, energy, material, services and seed expenses. In material inputs, we summarize fertilizer and plant protection. We use total livestock units per hectare as a proxy for manure. Except for land,

[4] Input data would be available at the farm level but information about the farm location would be available only at the federal state level due to data privacy reasons.
[5] We exclude the federal state *Saarland* and the cities of *Berlin*, *Hamburg* and *Bremen* due to their small geographical size and minor importance in wheat production.
[6] We preferred the TF-8 data given the higher representativeness; see SA for a comparison of TF-8 and TF-14 grouped data (Fig. S1).

labor and livestock, all inputs are deflated using national price indices provided by the German statistical agency (Statistisches Bundesamt, 2014) and normalized by the total utilized agricultural area per farm excluding fallow and set-aside land. We include an additional control variable for the share of spring wheat for each federal state and year from German official agricultural statistics (BMEL, 2015).

Land planted with wheat is considered to account for positive specialization and scale effects or negative yield effects resulting from marginal land (Kaufmann and Snell, 1997; Yang et al., 1992). On average, sample farms plant about 65 ha wheat. For historical reasons considerable differences in the farming structure (e.g., size, organizational and ownership structure, technology) between Eastern and Western Germany prevail (Tab. 2). To account for these we use a dummy variable indicating the five Eastern federal states.

Our data cover three major reforms of the EU's CAP that are known to provide incentives for planting crops (background information in the SA). To capture policy and other macroeconomic effects such as the price boom in 2008, we take time dummy variables into account.

2.4.2 Weather and phenological stages

We merge the annual FADN data with daily meteorological observations from 1218 weather stations and phenological data for winter wheat from 5671 stations scattered across Germany provided by the German Meteorological Service (Deutscher Wetterdienst, 2012; Deutscher Wetterdienst, 2014).

Tab. 1: Four defined phenological periods.

Period	Corresponding stages
1	Sowing – stem elongation (minus 1 day) [30]
2	Stem elongation – heading (minus 1 day) [51]
3	Heading – early milk ripening (minus 1 day) [73]
4	Early milk ripening – harvested product [99]

Note: stages following Meier (2001); decimal code in [ ].

For all federal states we distinguish four macro phenological periods (Tab. 1) and aggregate all weather variables accordingly (similar e.g., Butler and Huybers, 2015; details in SA). Temperature and solar radiation are mainly responsible for potential crop growth; however, day temperatures above the optimal level might induce heat stress and decrease wheat yields (Rötter and van de Geijn, 1999). To capture these effects, the average temperature is split into temperatures below and above an optimal temperature of 20 °C, above which growing conditions are likely not optimal (Rötter and van de Geijn, 1999). Accordingly, days with temperatures below the optimum but above the minimum of 4°C are denoted as growing degree days (GDD; expected positive effect on wheat growth). Temperatures above 20 °C, on the other, lead to heat stress and are summarized for each phenological period as killing degree days (KDD; expected negative impact on crop yields; Roberts et al., 2013, p. 237).

Crop water supply is determined by supply in the form of precipitation and atmospheric demand in the form of evapotranspiration. To appropriately account for water supply, we consider potential evapotranspiration according to Turc–Ivanov (ETP_{TI}) following Conradt et al. (2013) (see SA for details and an alternative measure according to Haude).

Marginal effects of additional water supply depend on actual levels and may switch signs. That is, precipitation might have a positive impact on plant growth if actual water supply is below a plant's optimum. On the other, precipitation might hamper growth of plants in case of water supply being greater than the plant's optimum. Since our weather variables are aggregated over the phenological periods, dry spells are not considered by the sole precipitation amount. Thus, in addition, we consider days without precipitation (DWP) to capture the distribution of precipitation (see SA for details).

Particularly in the eastern parts of Germany, low precipitation amounts might have a higher marginal effect than in the western parts, given the lower soil quality resulting from large shares of sand and low water-holding capacity. We account for this by an interaction of precipitation with the dummy variable $East_i$. Additionally, we consider solar radiation (SR in J cm^{-2}) and temperature normalized radiation (Gornott and Wechsung, 2016).

2.5 Econometric strategy

In what follows, we explain the two steps guiding our analysis.

2.5.1 Empirical model wheat yield variability in Germany 1995–2009

From an economic theorist's perspective, it seems important to consider variable and quasi-fixed inputs. However, economic theory has little to say about functional form, relationships among inputs, and the interrelation with weather variables (Coelli et al., 2005). Hence, our model building and selection approach is based on an empirical procedure suggested by Greene (2012, p. 178–80) to select variables with the objective of finding a suitable model that is robust against misspecification (further inspiration from Roberts et al., 2013). Given the rather short number of years available to us, we have to scrutinize a rich set of variables and focus on relevant information. For instance, the set of theoretically optimal weather variables at all phenological periods and the full potentially available set of inputs amounts to 18 and 8, respectively. This is not counting quadratic terms, interactions, alternative weather variables, and variations of functional form (log versus level).

First, we target at identifying an appropriate functional form for the production function relating output, wheat yield y_{it} of state i at period t, and inputs, denoted by x_{jit} where j indexes the inputs capital, labor, land used for wheat, energy, material, seed and manure. Second, the appropriate function relating yield and weather for each agronomic stage must be specified.

We consider wheat yield in log-differences ($\Delta \log y_{it}$) rather than absolute yields to approximate growth rates for two reasons. First, the log-ratio of wheat yields represents a relative change over time and is thus already a measure of relative variability. Second, analyzing first differences has advantages from a statistical point of view. Given the positive trends in the data (Fig. 1 for visual inspection), we account in a more flexible way for trends by means of the first differences compared to assuming a deterministic trend (Brown, 2013). Moreover, potential unit root problems (the Im, Pesaran and Shin test does not reject the null of a unit root at any conventional level) are usually resolved by first differencing (Chen et al., 2004). In addition, first-differencing eliminates unobserved heterogeneity effects likely present in panel data and reduces problems of serial correlation if data are persistent (Wooldridge, 2009).

Tab. 2: Summary statistics.

	Mean	sd	Min	Max
Wheat yield [100 kg ha^{-1}]	69.81	10.66	33.28	92.46
Per farm variables				
Land wheat [ha]	65.42	60.72	8.80	251.73
Land wheat East [ha]	126.80	47.06	46.45	251.73
Land wheat West [ha]	21.59	10.79	8.80	66.63
Total land [ha]	228.37	196.76	36.89	766.25
Total land East [ha]	448.30	94.72	196.90	766.25
Total land West [ha]	71.29	21.49	36.89	135.60
Capital [EUR ha^{-1}]	473.71	172.46	217.82	968.97
Labor [hours ha^{-1}]	51.31	23.06	22.45	142.63
Energy [EUR ha^{-1}]	171.25	41.75	106.36	270.85
Material inputs [EUR ha^{-1}]	308.96	63.03	163.92	517.41
Seeds [EUR ha^{-1}]	93.54	31.17	49.92	196.32
Manure [livestock units ha^{-1}]	0.37	0.20	0.06	0.91
Regional weather variables and controls				
Pot. evapotranspiration stage 1 [mm]	133.32	15.21	103.10	180.60
Prec. stage 1 [mm]	413.40	76.31	219.60	576.60
Prec. stage 1 East [mm]	413.00	78.67	227.2	562.30
Prec. stage 1 West [mm]	413.70	74.96	219.60	576.60
GDD stage 2 [°C]	324.74	33.50	212.96	422.26
Solar radiation stage 2 [kJ cm^{-2}]	60.22	6.49	37.71	77.05
Prec. stage 2 [mm]	74.50	20.80	26.00	117.02
KDD stage 3 [°C]	11.08	7.14	0.64	27.15
Prec. stage 4 [mm]	95.99	29.36	42.80	177.90
Prec. stage 4 East [mm]	97.28	30.91	42.80	176.30
Prec. stage 4 West [mm]	95.06	28.31	50.40	177.90
Share spring wheat [%]	1.97	1.49	0.30	8.59

Note: Data sources are Deutscher Wetterdienst (2012) and Deutscher Wetterdienst (2014), FADN (European Commission, 2015); share spring wheat: BMEL (2015).

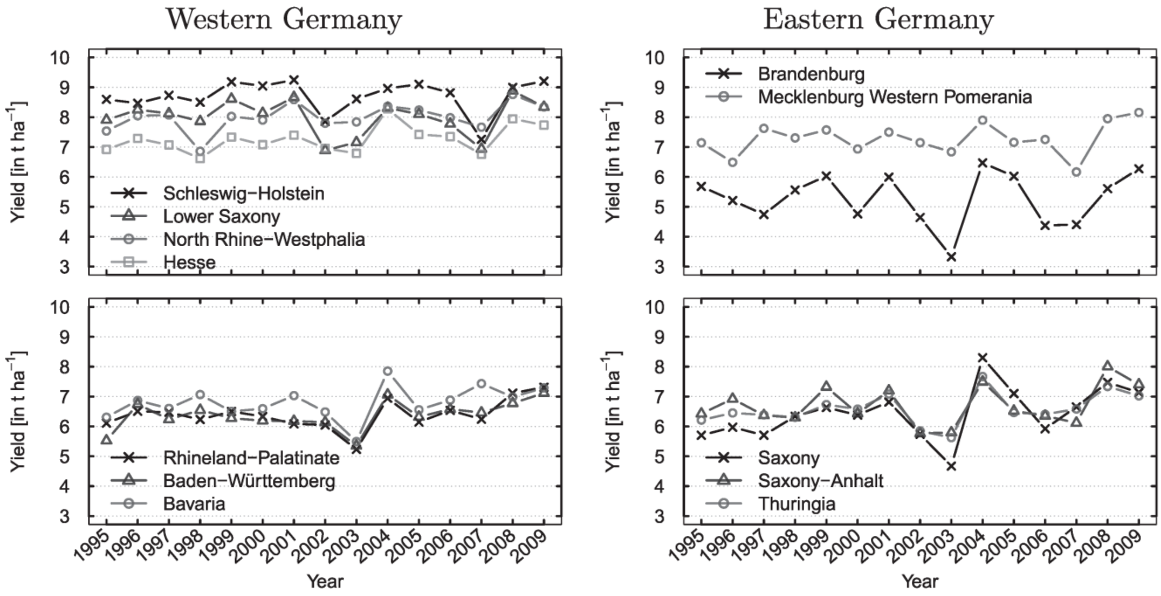


Fig. 1. Yields by federal state, 1995–2009.

This data transformation has the disadvantage of not directly quantifying the effect of, e.g., precipitation or technological change on yield and variance at the same time as proposed by Chen et al. (2004), based on the Just and Pope production function approach (see also Isik and Devadoss, 2006). However, log-differences still allow us to capture the effects of common yield shifts by the constant term, though not (directly) the impact of technological change. The effect measured by the constant accounts for several issues, including technological change, but also yield changes caused by the CO_2-fertilization effect (Attavanich and McCarl, 2014; see also Long et al., 2006, who discuss the potential size of the fertilization effect).

Accordingly, the logged wheat yield ratios are modelled as a function of three components: first, the production function with inputs, $f(x_{jit})$, where j indexes the inputs, and second, the weather function, $g(x_{kit})$, where k indexes the number of weather variables (each aggregated to four phenological sub-periods and counted as different variables).

The third component includes the number of other controls x_{sit} indexed by s and to isolate annual effects induced by policy changes, market, price and other shocks common to all federal states we include yearly dummy variables x_{st}. These dummy variables capture other shocks, including stochastic technological changes at the national level that depart from the linear federal state-specific trends and thus are not removed by first differencing. The annual dummy variables are also of importance from an econometric point of view since common cross-sectional dependence might prevail if such shocks are not addressed.

Our empirical strategy relies on a rich data set and by exploiting the panel data structure we eliminate potential time-constant sources of confounding by first-differencing. Simultaneous changes in the input-mix based upon expected yield variability (e.g., weather-induced) common to all federal states are captured by time dummy variables. Expectations as long as these result in land-adjustments are captured by the variable land. Also, marginal effects of inputs are estimated conditional on observed weather changes. Thus, if observed weather affects yield variability, this is accounted for while estimating these parameters.

Still, expectations about forthcoming yield changes specific to one or several (but not all) federal states might induce in-season-adjustments of material inputs such as fertilizer. Since these issues are usually unobservable, some inputs in our models might still be confounded with the error term. However, given that about 25% of the total agronomic management costs remain after seeding (KTBL, 2006), and could be adjusted based on very early yield forecasts during the growing season, we argue that the severity of such a simultaneity bias in our context falls within an acceptable range. Additionally, we address this problem by instrumental variables estimation in the robustness checks.

The base function is given by:

$$\Delta \log(y_{it}) = \Delta f(x_{jit}) + \Delta g(x_{kit}) + \Delta h(x_{st}, x_{sit}) + \Delta\varepsilon_{it} \tag{1}$$

herein $\Delta\varepsilon_{it}$ denotes the respective error term and symbol Δ indicates the first-differencing operator. Throughout the specification search, we work with yield in log-differences as dependent variable and all explanatory variables in first differences.

We basically test four models that share the same dependent variable but differ in the functional form for the production function $f(x_{jit})$and the weather component $g(x_{kit})$, as well as the included variables and interactions. Here we borrow from the idea of from general to simple modeling (Greene, 2012) but with the following rules to keep the number of parameters at a reasonable level (see Fig. 2 for an overview).

First, both considered production functions $f(x_{jit})$, that is, the transcendental logarithmic (translog) and quadratic including the full set of interactions, are simplified backwards and forwards to minimize the Akaike Information Criterion (AIC). We force linear terms into the model, while weather is left to the error term in step 1 (Fig. 2).

Second, both simplified functions are then merged with a theoretically optimal set of weather variables: GDD, precipitation, solar radiation, potential evapotranspiration according to Haude for all four stages, and KDD in stages 3 and 4. We consider two versions of the weather function $g(x_{kit})$: logs and levels (step 2a, Fig. 2). Accordingly, four possible combinations have to be tested: translog with weather either in logs or levels, and quadratic with weather either in logs or levels. The weather variables enter each of these models by phenological period. To select the relevant weather variables at the respective phenological stage, we rely on a backwards and forwards procedure to minimize the AIC (Roberts et al., 2013). Linear terms of weather variables were forced to be part of the model if quadratic terms were considered relevant. The remaining weather variables DWP, minimum and maximum temperatures, and potential evapotranspiration (Turc–Ivanov), are tested by phenological period and kept only if the AIC can be improved (step 2b).

Third, the resulting four models are enlarged as follows. Linear terms are added again if only squared terms and if only interactions are part of the model. Then, additional control variables as well as a full set of year dummy variables are tested (step 3a). Here we consider the share of spring wheat to account for possible lower yields if the share of spring wheat is high. Additionally, we account for the Oder river flood in 2002 and the European heat wave in 2003. Since extreme values might affect the point estimates, particularly those of the weather variables, we interact those federal states most affected by the flood (see SA) with the year dummy 2002 (x_6 , Eq. (4)). Since Brandenburg suffered most from the heatwave (Fig. 1), we interact this state with 2003 (x_7 , Eq. (4)). Again, we simplify the four enlarged models applying the back- and for- wards procedure to minimize the AIC (step 3b).

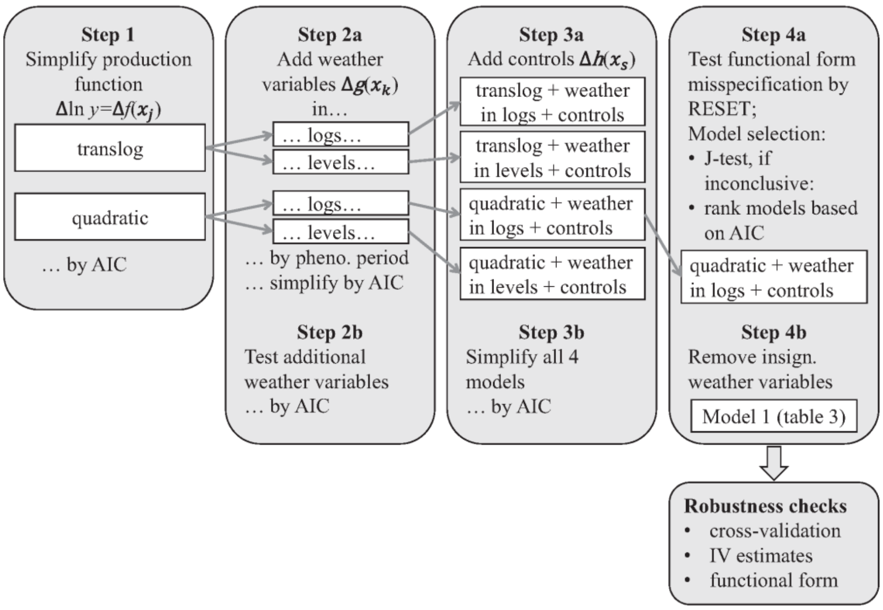


Fig. 2. Workflow of model building and selection. Note: AIC: Akaike Information Criterion; IV: instrumental variable; J-test: Davidson and MacKinnon J-test; pheno.: phenological; RESET: regression specification error test; stat. insig.: statistically insignificant (at 10% level).

Fourth, since we can rule out any remaining misspecification of functional form (RESET passed), we further carry out Davidson and MacKinnon J-tests to choose between these non-nested models. Since these results were inconclusive, we opted for the model with the lowest AIC (step 4a).[7] After all, the log-level specification, a commonly used approach in applied econometrics, e.g., Wooldridge (2009, p. 45–6), as well as in climate impact research (e.g., Lobell et al., 2011), seems to best fit to the input data. That is, the yields and the weather are modeled in logarithmic form (e.g., You et al., 2009), while inputs are modeled in levels. At the end of this procedure, we finally remove four more weather variables, which were statistically insignificant (robust standard errors; step 4b). Results were robust to inclusion, though we placed greater weight on econometric efficiency here (see SA, Tab. S8).

The final model, labeled as *model 1*, is given by Eq. (1) with $f(\cdot)$, $g(\cdot)$ and $h(\cdot)$ defined as follows:

$$f(x_{jit}) = \sum_{j=1}^{7} \beta_j x_{jit} + \frac{1}{2}\sum_{j=1}^{5} \beta_{jj}(x_{jit})^2 + \beta_{12}(x_{1it}x_{2it}) + \beta_{13}(x_{1it}x_{3it}) + \beta_{24}(x_{2it}x_{4it}) + \beta_{35}(x_{3it}x_{5it}). \tag{2}$$

Aside from services, all inputs x_{jit} remain in the final model specification. Capital (x_{1it}) appears in interaction with labor (x_{2it}) and seeds (x_{3it}), while labor with energy (x_{4it}), and seeds with manure (x_{5it}). Symbols x_{6it} and x_{7it} denote material inputs and wheat land, respectively. Seven weather variables in logarithmic terms enter the model in the spirit of a translog production function:

$$g(x_{kit}) = \sum_{k=1}^{7} \beta_k x_{kit} + \beta_{11}\frac{1}{2}\log(x_{1it})^2 + \sum_{k=2}^{3}\left[\beta_{k2}\log(x_{kit})\, East_i + \beta_{k3}\frac{1}{2}\log(x_{kit})^2\, East_i\right] \tag{3}$$

with x_{1it}: solar radiation stage 2, x_{2it}: precipitation stage 1, x_{3it}: precipitation stage 4, both interacted with a dummy variable for the Eastern German federal states $East_i$, x_{4it}: potential evapotranspiration ETP_{TI} stage 1, x_{5it} : GDD stage 2, x_{6it}: precipitation stage 2, x_{7it} : KDD stage 3.

$$h(x_{st}, x_{sit}) = \sum_{s=1}^{13} \beta_s x_{st} + \beta_{61}(x_6\, flood_i) + \beta_{72}(x_7\, Brandenburg_i) + \beta_{14}x_{14it} \tag{4}$$

herein x_1 to x_{13} denote annual dummy variables for 1997–2009 and x_{14it} the share of spring wheat on total wheat land. The dummy variable $flood_i$ is set to 1 for states that suffered from the flood in the year 2002; $Brandenburg_i$ denotes a dummy variable for this federal state.

[7] Tab. S6 in the SA shows the ranking of the four models that differ by their functional form of the right-hand side of the production function and the weather variables using the respective value of the AIC.

Given the interaction terms, we consider all variables except yields and dummy variables in mean-centered form, that is, each observation is normalized by its corresponding sample mean. Using panel data allows us to eliminate unobserved heterogeneity effects by first-differencing, and thus all models are estimated by Ordinary Least Squares (Greene, 2012). To address the potential endogeneity problem while estimating the effect of material inputs, we apply an instrumental variable (IV) estimation approach as a robustness check, where we use the second lagged differences of material inputs as instruments (see Tab. S10 in the supplementary appendix). The IV estimates reveal the same qualitative results (signs preserved) but a modestly higher estimate for material inputs; however, IV estimates are only consistent and do not epitomize an unbiased point of reference (Wooldridge, 2009, p. 510). Hence, we proceed upon the OLS estimation approach on first-differenced data, but provide the IV estimates in the supplementary appendix.[8]

To discuss the potential of an omitted variables bias, we present another model leaving out input variables similar to Miao et al. (2016, p. 201). To define *model 2*, within Eq. (1) all inputs (Eq. (2)) are removed. We assume that the specification search would have led to choosing the same weather variables.

2.5.2 Investigating the effect of inputs and weather on yield volatility

Model 1 passes the RESET procedure and hence we conjecture that model 1 is linearly separable in parameters (Tab. S7, SA). This is a pre-condition for an unbiased and reliable decomposition of the wheat yield variability as carried out in the second step of this empirical analysis.

Generally, two approaches to measure crop yield variability exist: absolute or relative. Chen et al. (2004), for instance, refer to an absolute measure. These authors rely in their rigorous analysis on a Just and Pope type production function approach to quantify weather effects on mean yield and its variance. In these type of models, however, the dependent variable must be stationary without first differencing, where our yield data require first differencing to ensure stationarity. Furthermore, absolute measures rely on absolute changes in yields, which might lead to seemingly increased risk if positive time trends prevail (Finger, 2010). Thus, we focus on relative risk measures to ensure the comparability of weather-induced wheat yield variability by region and over years. Relative variability, specifically volatility, can be measured for instance by using the standard deviation of time differences of logged yields (log-returns), or by a coefficient of variation (CV) for a time series of yield levels (Finger, 2010, p. 177; Ray et al., 2015, p. 2).

In this second step, we extract and analyze weather-determined yield variability, similar to Osborne

[8] Dynamic modeling approaches fail in this context due to the limited number of observations.

and Wheeler (2013). We augment these authors' approach by adding inputs, a wider range of weather variables, and analyzing a regional panel data set inspired by You et al. (2009, p. 1012). In contrast to Osborne and Wheeler (2013, p. 4, 7) and Ray et al. (2015, p. 2), we separate the weather explained portion of the yield from the input-determined part (see Fig. 3). In addition, our volatility measure directly refers to the weather-induced variation of yields, while the measure "*climate explained yield variability*" of Ray et al. (2015, p. 2) is in absolute terms not independent of the yield dimension.

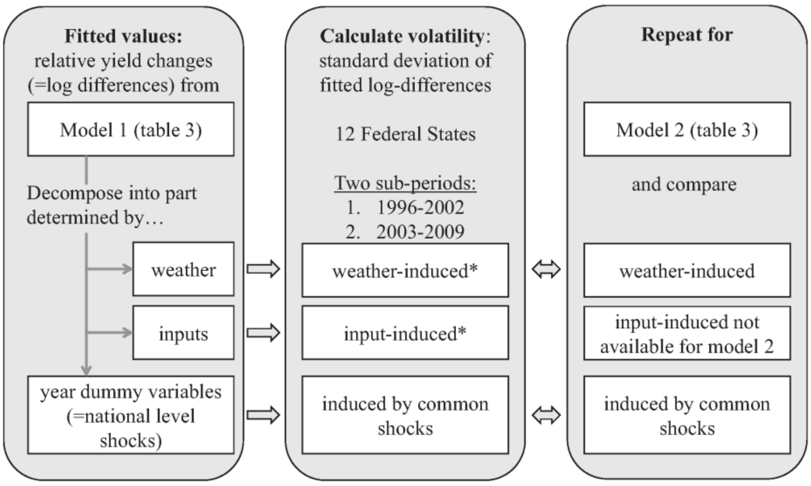


Fig. 3. Workflow of volatility decomposition. *: decomposition illustrated in Fig. 4. Full results in SA, Tab. S11.

Weather caused relative yield changes $\widehat{y_{it}^{W}}$ are defined as the fitted first differences of log yields in the regression model (Eq. (1)) resulting from weather variations following Eq. (3). That is, inputs are evaluated at their means (zero since in mean-centered form) and other controls such as year dummy variables are not included: $\widehat{y_{it}^{W}} = \Delta \widehat{\log(y_{it})} = \Delta \widehat{g(y_{it})}$. Based on the fitted series we extract weather-induced volatility $\widehat{v_i^W}$ for each state $i = 1, \dots, n$ using the standard deviation over $\widehat{y_{it}^{W}}$:

$$\widehat{v_i^W} = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}\left(\widehat{y_{it}^{W}} - \overline{\widehat{y_i^W}}\right)^2} \tag{5}$$

with year $t = 1, \dots, T, T = 7$ and mean $\overline{\widehat{y_i^W}} = \frac{1}{T}\sum_{t=1}^{T}\widehat{y_{it}^{W}}$.

To capture changes in weather-induced risk over time, we divide the sample into two equal-length sub-periods: 1996–2002 and 2003–2009. To compare weather-induced volatility changes and those induced by inputs, we adopt this approach for input-determined yield changes accordingly. For additional comparisons, we calculate volatilities based on fitted values allowing both – inputs and weather – to fluctuate while excluding controls and year dummy variables.

2.6 Results and discussion

First, we report the results of the production function estimation and the robustness checks; second, we discuss the estimated volatilities as determined by weather and inputs. In Tab. 3 we present the estimates for two models.[9] While model 1 refers to the fully specified model, model 2 is dedicated to the omitted- variables-bias discussion. All inputs are in levels. Hence, the coefficients can be interpreted as semi-elasticities while the weather variables enter in logs; these estimates can be interpreted as elasticities. A Davidson and MacKinnon J-test and AIC values show that model 1 is superior to a log–log model (see SA, Tab. S9, model 5).

2.6.1 Production function inputs

The considerable number of statistically significant interaction terms underpins our choice of a flexible functional form. *Energy* and *material* inputs reveal positive linear and negative quadratic effects, though energy only does so in a significant quadratic term. That is, material inputs positively affect yield changes with decreasing marginal productivity. Starting from the sample mean, a 10% increase (roughly 30 Euros per hectare) leads to positive yield changes of about 1.7%. Because of the non-linear relationship, following a 30% increase (roughly 90 Euros), yields would already decrease by 1.06%.

Seeds show a negative linear and positive quadratic coefficient. This implies that, starting from the sample mean, a reduction of seeds would lead to soaring yields, whereas increases would first lead to decreases and then to increases again. This effect might be traced back to the variable definition in monetary terms. Reducing seeds would lead to increasing yields, though low seed densities require nearly perfect water supply conditions. Agronomic relations might explain the negative range. Late sowing requires a higher rate of seed input per hectare to ensure the full development of a plant population. On the other hand, too dense populations can reduce yields.

[9] The p-values are based on spatial correlation consistent (SCC) standard errors, which are also robust to (cross-) serial correlation (Driscoll and Kraay, 1998; Millo, 2014). Note that the time dimension of our data is relatively small (T = 14) and as such at the lower bound (Driscoll and Kraay, 1998, p. 556).

Tab. 3: Effects of inputs and weather on relative wheat yield variability in Germany, 1996–2009.

	(1) Final specification	(2) Drop inputs
Intercept	0.072 (0.0116)***	0.083 (0.0127)***
Inputs		
Capital	0.043 (0.0399)	
Labor	-0.195 (0.1743)	
Energy	0.123 (0.0765)	
Material inputs	0.091 (0.0305)***	
Seeds	-0.142 (0.0495)***	
Manure	-5.540 (6.9399)	
Land wheat	-0.050 (0.0418)	
Capital squared	0.001 (0.0002)***	
Labor squared	0.016 (0.0095)*	
Energy squared	-0.011 (0.0020)***	
Material inputs squared	-0.001 (0.0002) ***	
Seeds squared	0.004 (0.0011)***	
Capital · labor	-0.004 (0.0015)***	
Capital · seeds	-0.001 (0.0001)***	
Labor · energy	0.011 (0.0036)***	
Seeds · manure	0.288 (0.1611)*	
Weather		
Prec. stage 1	0.050 (0.0159)***	0.022 (0.0194)
Prec. stage 1 · East	-0.167 (0.0439)***	-0.139 (0.0783)*
Prec. stage 1 squared · East	-0.913 (0.1412)***	-0.970 (0.3244)***
Pot. evapotranspiration stage 1	0.225 (0.0512)***	0.250 (0.0757)***
Growing degree days stage 2	0.114 (0.0313)***	0.125 (0.0727)*
Solar radiation stage 2	-0.023 (0.0423)	0.034 (0.0743)
Solar radiation stage 2 squared	0.442 (0.1561)***	0.133 (0.1236)
Prec. stage 2	-0.019 (0.0067)***	-0.015 (0.0091)*
Killing degree days stage 3	-0.018 (0.0027)***	-0.017 (0.0040)***
Prec. stage 4	-0.039 (0.0164)**	-0.044 (0.0195)**
Prec. stage 4 · East	0.084 (0.0519)	0.095 (0.0408)**
Prec. stage 4 squared · East	0.262 (0.1615)	0.261 (0.1064)**
Controls		
Share land spring wheat	-1.020 (0.4498)**	-0.840 (0.6020)
Year 2003 · Brandenburg	-0.388 (0.0229)***	-0.386 (0.0145)***
Flood 2002	-0.073 (0.0134)***	-0.103 (0.0152)***
Year 1997	0.005 (0.0189)	-0.058 (0.0071)***
Year 1998	-0.051 (0.0283)*	-0.149 (0.0192)***
Year 1999	-0.073 (0.0494)	-0.180 (0.0438)***
Year 2000	-0.199 (0.0485)***	-0.334 (0.0466)***
Year 2001	-0.217 (0.0584)***	-0.340 (0.0555)***
Year 2002	-0.311 (0.0677)***	-0.476 (0.0670)***
Year 2003	-0.466 (0.0860)***	-0.673 (0.0892)***
Year 2004	-0.351 (0.0860)***	-0.533 (0.0973)***
Year 2005	-0.511 (0.1075)***	-0.755 (0.1220)***
Year 2006	-0.602 (0.1175)***	-0.856 (0.1294)***
Year 2007	-0.722 (0.1196)***	-0.946 (0.1338)***
Year 2008	-0.638 (0.1349)***	-0.889 (0.1512)***
Year 2009	-0.688 (0.1482)***	-0.994 (0.1681)***
R^2	0.83	0.74
Adj. R^2	0.77	0.69

Note: Dependent variable: first differences of logged wheat yield. Num. obs.: 168. Weather in logs, inputs in levels. Coefficients/standard errors for inputs multiplied by 100. Explanatory variables first differenced; weather/inputs mean centered. Spatial and serial correlation robust standard errors in () (Driscoll and Kraay, 1998). Pot.: potential; prec.: precipitation. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Turning to *capital* and *labor*, two quasi-fixed inputs, we find positive linear and quadratic effects of capital changes on yield growth rates but negative interaction effects with labor (and seeds). For labor, we find negative linear and positive quadratic terms. That is, any deviation from the sample mean in labor would cause increases in the yield growth rate evaluated at the sample mean. These estimates should, however, be interpreted in light of the observed trend in the data to reduce overall capital and labor input per hectare from 1999 on. To disentangle the capital-labor relationship, we follow the idea of simple slopes in two-way interactions and use capital as a moderator. We find for *low levels of capital* (mean minus standard deviation) that a reduction of labor input negatively affects yield changes, whereas additional labor units at low capital levels contribute positively to the growth rate of wheat yields. For *high capital levels* (mean plus standard deviation), the substitution effect is less obvious: deviating from the sample mean level of labor causes positive impacts, whereas the positive effect of reducing labor input is more pronounced. This implies that in high capital production systems, capital productivity might be improved through labor reduction.

The effect for *land planted with wheat* remains insignificant. This might be traced back to two opposing effects. First, due to specialization and scale effects, we would expect yield to increase in land planted with wheat. Second, more marginal land might be used for cropping wheat in the course of time, incentivized for instance by rising wheat prices (Haile et al., 2016). In addition, such land might be more sensitive to changes in weather conditions, and thus we would expect a negative effect of a higher share of marginal land on wheat yields. While the first effect might be particularly relevant for Western Germany, the second effect might be more relevant for the eastern part.

All time dummy variables are significant from 2000 onwards; all are negative. That is, compared to changes between 1995/96, the wheat growth rate decreases for these years. Changes in the relative competitiveness of wheat due to decoupled direct payments within the CAP (starting from 2000 onwards) might provide an explanation for this finding. These coefficients capture common yield shifts, also possibly due to weather or other common macroeconomic shocks (e.g., prices, technological change), the effects of which cannot be isolated. For example, our modeling approach does not allow us to identify the direct impact of technological change, nor to disentangle the effect of CO_2 on yield variability as shown by Attavanich and McCarl (2014). The time dummy variables for the flood and the European heatwave reveal significant negative impacts on the wheat growth rates for the respective federal states.

2.6.2 Weather

In the first early development stage of the plant (sowing/end of tillering), we find significant positive effects of precipitation levels and potential evapotranspiration (ETP_{TI}): a 10% increase in ETP_{TI} would lead to a 2.3% increase in yields. A sufficient water supply improves biomass production, determining the yield potential of the plant (Chmielewski and Köhn, 2000). Also, Roberts et al. (2013) find positive effects of a vapor pressure deficit, a main component of the evapotranspiration measure. For the Eastern German federal states, which are known for soil conditions with low water holding capacity, we find negative coefficients for the linear and quadratic terms of precipitation in stage 1. Given the predominating soil condition in Eastern Germany, nutrient leaching might be problematic at higher precipitation levels in the first phenological stage. This might hamper yield potentials. For the fourth stage (early milk ripening/harvesting), we find negative impacts of precipitation, that is, as expected, in late-ripening crops, additional water may lead to yield losses. However, for Eastern Germany, we find positive impacts of precipitation; marginally significant with $p = 0.106$ (linear term) and $p = 0.108$ (squared term). Since the fourth stage includes early milk ripening, in which water scarcity is very likely in the Eastern federal states compared to other regions in Germany, additional water supply could foster increased yield quantities (Fricke and Riedel, 2015).

For the second stage (stem elongation until heading), we find negative effects of precipitation and positive effects for GDD. That is, temperature below the optimal level positively influence wheat growth and thus yields. Solar radiation has a non-linear positive effect that is attributable to increased photosynthesis (Roberts et al., 2013). An increasing water supply in this developmental stage, however, rather hinders growth as indicated by the negative coefficient. This result points to water supply being close to optimum on average. In the third stage (heading/ early milk ripening), we find considerable impacts of temperature: KDD affect yield negatively, which is in line with existing research (Roberts et al., 2013). These effects, however, remain small: starting from the sample mean, an increase in KDD by one standard deviation (approx. 65%) would lead to yield losses of 1.2%.

Despite taking regional KDD into consideration, we find an additional significant effect of the heatwave for Brandenburg. This might be traced back to Brandenburgs natural conditions, particularly sandy soils with low water-holding capacity (Wessolek and Asseng, 2006) and uncaptured soil-specific heatwave dynamics (Perkins, 2015). Together with different effects of precipitation for Eastern and Western Germany in two phenological stages, our results reveal the importance of the spatial-temporal distribution of water supply and its dependence on soil conditions.

To summarize, all weather effects can be grounded on agronomic-theoretical explanations and are in line with previous findings. Our variable selection and phenological data aggregation reflect the complexity of yield formation. Model 1 may be criticized regarding the inclusion of the extreme weather years. Thus, we performed a leave-one-out cross-validation confirming the robustness of the model (e.g., Blanc, 2012; details in SA).

2.6.3 Decomposing wheat yield volatility

To answer the question of how volatility differs across regions and over time, as well as to disentangle its drivers, we decompose the standard deviation of the wheat growth rates (see Fig. 3). We illustrate these measures in Fig. 4 (based on Tab. S11, SA). Actual, weather- and input-induced volatilities are plotted for two sub- periods: 1996–2002 and 2003–2009 (grey-solid and grey-dashed circles). Averaging over regions and time, inputs explain ca. 49% of the total actual wheat yield volatility, while weather explains 43% (evaluated at the sample means, based on values in Tab. S11, SA). Comparing actual volatilities for the sub-periods over time, wheat yield volatility increases except for one state (North Rhine-Westphalia, Fig. 4-a). Riskier areas regarding weather and inputs are found in the eastern part of Germany.

We use regional aggregated yield data at the federal state level. Spatially uncorrelated risks, that is, idiosyncratic shocks, "*self-diversify*" at this higher aggregation level compared to firm-level data, while more systemic variation remains (Woodard and Garcia, 2008, p. 37; Marra and Schurle, 1994).[10] Hence, weather- induced volatility at the state level can be interpreted as a measure of systemic weather risk in agricultural production (cf. Xu et al., 2010, p. 267–268). As illustrated in Fig. 4-b, weather-caused volatility differs slightly by region with higher volatilities in the eastern part. Comparing these volatilities with those caused by input adjustments (Fig. 4-c), we find for the entire eastern region, as well as some western regions, higher input-induced volatilities compared to the volatilities traced back to weather changes (e.g., Bavaria). Over time, we observe increases in actual volatility, on average. However, this can only be traced back to joint increases in weather and inputs in some regions (e.g., Saxony), while in other regions weather- and input-induced volatility changes reveal opposite signs. For instance in Brandenburg, weather-induced yield volatility considerably increases but input-induced volatility decreases. Still, the overall increase of actual volatility cannot be fully traced back to weather and input adjustments.

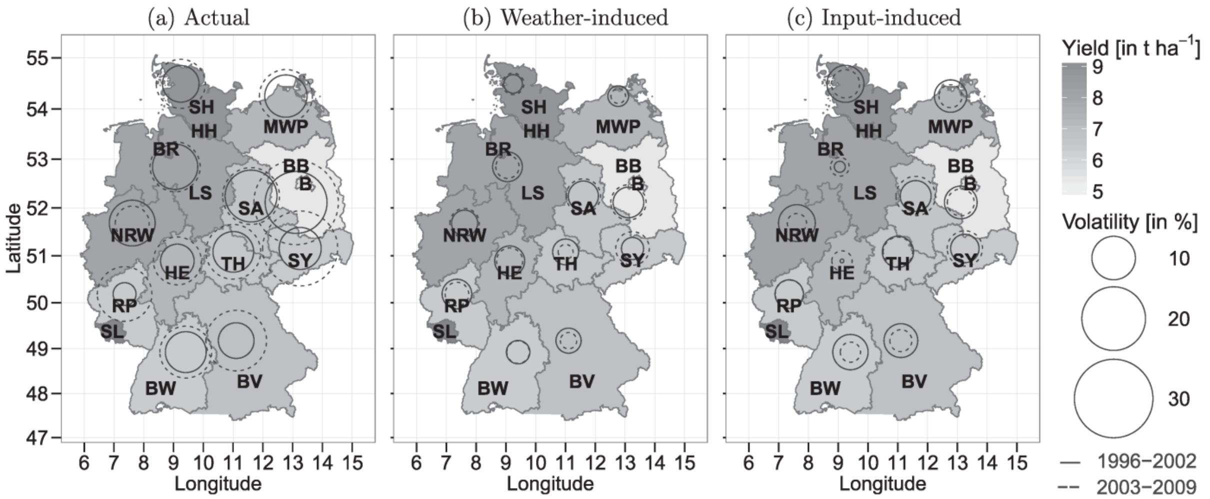


Fig. 4. Actual, weather- and input-induced wheat yield volatility for two sub-periods. Note: Bubbles indicate volatilities of different magnitudes. B: Berlin, BB: Brandenburg, BR: Bremen, BV: Bavaria, BW: Baden-Wuerttemberg, HE: Hesse, HH: Hamburg, LS: Lower Saxony, MWP: Mecklenburg-West Pomerania, NRW: North Rhine-Westphalia, RP: Rhineland-Palatinate, SA: Saxony-Anhalt, SH: Schleswig–Holstein, SL: Saarland, SY: Saxony, TH: Thuringia.

[10] Note that the illustration of the aggregation argument by Woodard and Garcia (2008, p. 37–38) does not acknowledge that a part of the weather risk might be included among the idiosyncratic risks that self-diversify. Precipitation and related variables that are functions of the latter are expected to vary more across space than temperature.

We find a higher share of explained actual volatility, for the first compared to the second period. To illustrate an extreme, in Hesse from 1996–2002, about 93% of the actual volatility are explained (6.11 *volatility inputs and weather;* 6.55 *actual volatility*, values according to Tab. S11, SA); from 2003–2009 this amounts to 35%, however. Averaging over all regions, we find 83% of the actual volatility explained by inputs and weather from 1996–2002, while 44% in the period 2003–2009. These findings can be explained in part by the use of time dummy variables in the regression model, which are isolated in the volatility measures for inputs and weather but are particularly important in the second period. Hence, we likely underestimate the weather effect because common weather shocks are captured by time dummy variables. As a robustness check, we investigated whether the different geographical sizes of the states affect our results (aggregation bias); this is not the case (see SA for details).

Weather-driven volatility at the state level seems to be rather low given that we would expect higher changes caused by varying weather conditions. In other words, we conjecture that systemic risk cannot only be traced back to weather as measured in our model (regional temperature, solar radiation, precipitation and evapotranspiration). Common shocks at the macro (i.e. national) level are relevant. The latter include weather extremes but also policy and price changes affecting many farms as well as consequential input-adjustments. The significant year dummy variables from 2000 onwards capture exactly such macroeconomic and policy changes (Tab. 3). In this period several reforms of the CAP affected farms production (intensity) decisions, for instance, the de-coupling of direct payments from the crop being planted starting in Germany in 2000, which was discussed in 2003, reinforced in 2005 and verified in 2008. The price boom for agricultural commodities in 2007/08 also occurred during this period. The national level volatility based on year dummy variables reflects the increasing importance of common shocks: 5% in 1996–2002, 11.2% in 2003–2009 (see Fig. 3 and Tab. S11, SA).

How would the results, particularly the weather-induced volatility, look like if input adjustments were neglected? Similar to the full model 1, the estimates for the reduced form model 2 reveal increases in weather-induced volatility for some regions, while for others decreasing measures for the second period prevail (Tab. S11, SA). The unexplained share of volatility and the national level volatility are also higher in the second period. Averaging over regions and time, weather explains practically the same fraction of the total actual volatility in model 2 (both models: 43.5%). Comparing the regional volatility estimates within period 1996–2002, some are overestimated in model 2 (e.g., Saxony-Anhalt and Brandenburg) while some are underestimated (e.g., Hesse). In period 2003–2009, the majority of the measures of the reduced form model overestimate the weather component, though in some regions only by a minor rate. For only one state (Hesse) the two models differ qualitatively: while the full model detects decreases in weather-induced volatility in period 2003–2009, model 2 finds small increases. Thus, one could draw misleading conclusions regarding the weather- induced volatility in both sign and size while neglecting inputs.

However, and most importantly, the unexplained part is higher compared to the full model. As a consequence, too much emphasis would be placed on the interpretation of the common shocks (significant for all years in model 2, considerably higher estimates; this results in higher national level volatility). As such, the systemic macro risk would be overestimated. At the same time, input- adjustments as possible consequences of price and policy shocks, which are simply rational adaptation by farmers, would not be discussed at all.

2.7 Concluding remarks

Wheat is a major commodity that plays a crucial role for food security. The recently observed increase in relative wheat yield variability for Germany – an important wheat producer in the EU – begs the question: Can these increases be traced back to weather changes? Or is it “simply” the result of farms’ adaptations to changing institutional and macroeconomic conditions leading to adjustments in their input-mix? To answer these questions, we analyze relative wheat yield variability consistent with production economics and agronomic climate impact research. We use a rich set of regional accountancy data and weather variables at the respective phenological stages from 1996 to 2009. Obtained wheat yield volatilities are decomposed into weather- and input- driven categories.

In line with production economics and agronomic research, we find that both inputs and weather impact relative yield changes. Common shocks at the national level play a significant role from 2000 onwards, a period characterized by fundamental changes in the EU’s CAP and price booms for agricultural commodities. Decomposing wheat yield volatility reveals regionally heterogeneous weather-induced instabilities. Splitting the sample into two sub-periods, we find increases in actual volatility over time, where macro-level shocks including weather extremes contribute. These increases, however, can only in some regions be traced back to joint increases of the weather-induced component and the part caused by adjustments in the input-mix. A number of regions even show decreases in weather-caused volatility over time.

This study is relevant for several reasons. First, future climate impact analyses, which inform policy makers, could utilize this case study as a proof of concept. We could show that omitting inputs would rarely alter our results in a qualitative manner, though would do so quantitatively. Weather impacts and common shocks would be overestimated in the case of leaving out input choices, and adjustments in the input mix would not be discussed at all. We thus contribute to the debate of whether inputs should be a part of climate impact research, where purely statistical approaches are still prominent (Liu et al., 2016; Miao et al., 2016). To conclude, independent of the model type, relating yield and weather offers reasonable results and valid approximations.

Second, these insights support approaches such as the European Commission's MARS[11] project, which is relevant for policy makers for crisis intervention. Considering yield vulnerability by phenological stage at the regional level could improve the seasonal forecasting of potential crop shortages attributable to weather. Better knowledge about yield vulnerability might also help farmers adjust their agronomic management to better cope with downside risk (Chipanshi et al., 2015).

Third, wheat yield vulnerability by phenological stage at the regional level might also be of interest for insurance design and modeling weather risk (Conradt et al., 2015; Odening and Shen, 2014). Since our approach decomposes the influences of weather- and input-related impacts on wheat yields and averages out idiosyncratic shocks, it might help insurers to improve the determination of insurance claims. For insurers, it might be relevant to only indemnify weather-related yield losses. In addition, a more cost-effective assessment of common weather- related yield losses might enable insurance companies to better cope with systemic risk. This would benefit both the insurer and the insured (cf. Finger, 2013). Insurance-based solutions have recently gained attention because of their potential to contribute to stabilizing farm incomes and thus food security, particularly in regions where smallholder farming prevails (Surminski et al., 2016).

Since many of the European CAP reforms aim to reduce the impact that disbursed subsidies have on input-intensity and to protect the environment at the same time (Levers et al., 2016), our results may further offer insights into how wheat yield variability is related to the interplay between weather, input-use and agricultural policy. These may help investigate how to reduce distorting policy impacts on input-intensity, while taking into account that these choices also relate to farmers' risk mitigation strategies. This in turn is a pre-condition for ensuring secure, resilient and sustainable food production. The recently-established risk management toolkit under pillar two measures of the EUs CAP might offer a reasonable starting point, though it has seen heterogeneous acceptance among member states thus far. Additionally, as shown by Gaupp et al. (2016), wheat yields are independent at the global level. This is a pre-condition to stabilize food supply by international trade, which could be another option for policy makers (e.g., Brown et al., 2017).

The work presented here displays some shortcomings. First, our data do not allow us to account for land use changes. For instance, farmers might reallocate land for highly subsidized renewable energy plants closer to the farm to save transportation costs. As a consequence, wheat might be reallocated to more distant plots, possibly with lower soil quality. The land used for wheat would not change overall, but yields would be more sensitive to weather impacts. Additionally, yield variability might be subject to technological progress (Chen et al., 2004), which could not explicitly be modelled with our data set. Using improved technology to increase agricultural productivity could be one way to

[11] Monitoring Agricultural Resources: https://ec.europa.eu/jrc/en/mars; accessed 16.02.2017.

globally ensure the stable supply of sufficient food quantities (Pardey et al., 2016). As shown by Emerick et al. (2016), against the backdrop of increasing weather risk, particularly new seed varieties with a reduced downside risk have the potential to crowd-in inputs such as fertilizer to increase yields (in addition to the positive agronomic effect on yield). Given that farmland expansion is already at its limit and in some regions only possible at the costs of biodiversity (e.g., Foley et al., 2011), such technical change could contribute to closing yield gaps with less negative environmental impact that the pure intensification of crop production by increasing fertilizers or irrigation would likely have.

Finally, from a producer’s perspective, economic risk matters as well. Output price variation has increased in the recent decade and proven to reduce production intensity (Haile et al., 2016). Future research analyzing weather impacts on agricultural production should thus consider farm-level input adaptations to changes in weather- and price-risk as well as policy changes and macroeconomic developments.

2.8 References

Attavanich, W., McCarl, B.A., 2014. How is CO_2 affecting yields and technological progress? A statistical analysis. Clim. Chang. 124 (4), 747–762.

Banse, M., van Meijl, H., Tabeau, A., Woljer, G., 2008. Will EU biofuel policies affect global agricultural markets? Europ. Rev. Agr. Econ. 35 (2), 117–141.

Barkaszi, L., Keszthelyi, S., Csatári, E., Pesti, C., 2009. FADN accountancy framework and cost definitions. In: FACEPA Deliverable. No. D1.1.1. http://facepa.slu. se/documents/Deliverable_D1-1-1_LEI.pdf (Accessed 24.03.2016).

Blanc, E., 2012. The impact of climate change on crop yields in Sub-Saharan Africa.

Am. J. Clim. Chang. 1 (1), 1–13.

BMEL, 2015. Statistik und Berichte des Bundesministerium für Ernährung und Landwirtschaft. http://www.bmel-statistik.de/ (Accessed 23.03.2016).

Brown, I., 2013. Influence of seasonal weather and climate variability on crop yields in Scotland. Int. J. Biometeorol. 57 (4), 605–614.

Brown, M., Antle, J., Backlund, P., Carr, E., Easterling, W., Walsh, M., Ammann, C., Attavanich, W., Barrett, C., Bellemare, M., Dancheck, V., Funk, C., Grace, K., Ingram, J., Jiang, H., Maletta, H., Mata, T., Murray, A., Ngugi, M., Ojima, D., O'Neill, B., Tebaldi, C., 2015. Climate Change, Global Food Security and the US Food System. US Global Change Research Program. http://www.usda. gov/oce/climate_change/FoodSecurity2015Assessment/FullAssessment.pdf (Accessed 06.02.2017).

Brown, M.E., Carr, E.R., Grace, K.L., Wiebe, K., Funk, C.C., Attavanich, W., Backlund, P., Buja, L., 2017. Do markets and trade help or hurt the global food system adapt to climate change? Food Policy 68, 154–159.

Butler, E.E., Huybers, P., 2015. Variations in the sensitivity of US maize yield to extreme temperatures by region and growth phase. Environ. Res. Lett. 10 (3), 1–8.

Calderini, D.F., Slafer, G.A., 1998. Changes in yield and yield stability in wheat during the 20th century. Field Crop. Res. 57, 335–347.

Chen, C.-C., McCarl, B.A., Schimmelpfennig, D.E., 2004. Yield variability as influenced by climate: a statistical investigation. Clim. Chang. 66 (1-2), 239–261.

Chipanshi, A., Zhang, Y., Kouadio, L., Newlands, N., Davidson, A., Hill, H., Warren, R., Qian, B., Daneshfar, B., Bedard, F., Reichert, G., 2015. Evaluation of the integrated Canadian crop yield forecaster (iccyf) model for in-season prediction of crop yield across the Canadian agricultural landscape. Agric. Forest Meteorol. 206, 137–150. .

Chmielewski, F.-M., Köhn, W., 2000. Impact of weather on yield components of winter rye over 30 years. Agric. For. Meteorol. 102, 253–261.

Coelli, T.J., Rao, D.S.P., O'Donnell, C.J., Battese, G.E., 2005. An Introduction to Efficiency and Productivity Analysis, second ed., Springer, New York.

Conradt, S., Finger, R., Bokusheva, R., 2015. Tailored to the extremes: quantile regression for index-based insurance contract design. Agr. Econ. 46 (4), 537– 547.

Conradt, T., Wechsung, F., Bronstert, A., 2013. Three perceptions of the evapotranspiration landscape: comparing spatial patterns from a distributed hydrological model, remotely sensed surface temperatures, and sub-basin water balances. Hydrol. Earth Syst. Sci. 17 (7), 2947–2966.

Dell, M., Jones, B.F., Olken, B.A., 2014. What do we learn from the weather? The new climate-economy literature. J. Econ. Lit. 52 (3), 740–798.

Deschênes, O., Greenstone, M., 2007. The economic impacts of climate change: evidence from agricultural output and random fluctuations in wheather. Am. Econ. Rev. 94 (1), 354–385.

Deutscher Wetterdienst, 2012. Daten der Klimastationen des Deutschen Wetterdienstes. Offenbach.

Deutscher Wetterdienst, 2014. Phänologie – Daten Deutschland. http://www.dwd. de/DE/klimaumwelt/klimaueberwachung/phaenologie/daten_deutschland/ daten_deutschland_node.html (Accessed 01.12.2015).

Driscoll, J.C., Kraay, A.C., 1998. Consistent covariance matrix estimation with spatially dependent panel data. Rev. Econ. Statist. 80, 549–560.

Emerick, K., de Janvry, A., Sadoulet, E., Dar, M.H., 2016. Technological innovations, downside risk, and the modernization of agriculture. Am. Econ. Rev. 106 (6), 1537–1561.

European Commission, 2015. FADN – Public Database. http://ec.europa.eu/ agriculture/rica/database/consult_std_reports_en.cfm (Accessed 30.11.2015).

FAOstat, 2015. Production World, European Union, and Germany 1995–2013. http://faostat3.fao.org/download/Q/QC/E (Accessed 30.11.2015).

Finger, R., 2010. Evidence of slowing yield growth – the example of Swiss cereal yields. Food Pol. 35, 175–182.

Finger, R., 2013. Investigating the performance of different estimation techniques for crop yield data analysis in crop insurance applications. Agr. Econ. 44 (2), 217–230.

Foley, J.A., Ramankutty, N., Brauman, K.A., Cassidy, E.S., Gerber, J.S., Johnston, M., Mueller, N.D., O'Connell, C., Ray, D.K., West, P.C., Balzer, C., Bennett, E.M., Carpenter, S.R., Hill, J., Monfreda, C., Polasky, S., Rockström, J., Sheehan, J., Siebert, S., Tilman, D., Zaks, D.P.M., 2011. Solutions for a cultivated planet. Nature 478 (7369), 337–342.

Fricke, E., Riedel, A., 2015. Raps und Getreide nicht zu früh beregnen. www.lwk- niedersachsen.de/download.cfm/file/185,28fd973a-a725-05eb- 3a7cd238ee115478pdf.html (Accessed 23.03.2016).

Gaupp, F., Pflug, G., Hochrainer-Stigler, S., Hall, J., Dadson, S., 2016. Dependency of crop production between global breadbaskets: a copula approach for the assessment of global and regional risk pools. Risk Anal. Early View, 1–17.

Gohin, A., 2006. Assessing cap reform: sensitivity of modelling decoupled policies. J. Agr. Econ. 57 (3), 415–440.

Gornott, C., Wechsung, F., 2016. Statistical regression models for assessing climate impacts on crop yields: a validation study for winter wheat and silage maize in Germany. Agric. For. Meteorol. 217, 89–100.

Greene, W.H., 2012. Econometric Analysis, seventh ed. Pearson, Boston, Mass, international Edition.

Haile, M.G., Kalkuhl, M., von Braun, J., 2016. Worldwide acreage and yield response to international price change and volatility: a dynamic panel data analysis for wheat, rice, corn, and soybeans. Amer. J. Agr. Econ. 98 (1), 172–190.

Iglesias, A., Quiroga, S., 2007. Measuring the risk of climate variability to cereal production at five sites in Spain. Clim. Res. 34 (1), 47–57.

IPCC, 2014. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (Eds.)]. IPCC, Geneva.

Isik, M., Devadoss, S., 2006. An analysis of the impact of climate change on crop yields and yield variability. Appl. Econ. 38 (7), 835–844.

Kaufmann, R.K., Snell, S.E., 1997. A biophysical model of corn yield: Integrating climatic and social determinants. Am. J. Agr. Econ. 79 (1), 178–190.

Krause, J., 2008. A Bayesian approach to German agricultural yield expectations.

Agr. Finance Rev. 68 (1), 9–23.

KTBL, 2006. Betriebswirtschaftsplanung Landwirtschaft 2006/07, 20th ed. KTBL (Kuratorium für Technik und Bauwesen in der Landwirtschaft), Darmstadt.

Levers, C., Butsic, V., Verburg, P.H., Müller, D., Kuemmerle, T., 2016. Drivers of changes in agricultural intensity in Europe. Land Use Policy 58, 380–393.

Liu, B., Asseng, S., Müller, C., Ewert, F., Elliott, J., Lobell, D.B., Martre, P., Ruane, A.C., Wallach, D., Jones, J.W., Rosenzweig, C., Aggarwal, P.K., Alderman, P.D., Anothai, J., Basso, B., Biernath, C., Cammarano, D., Challinor, A., Deryng, D., Sanctis, G.D., Doltra, J., Fereres, E., Folberth, C., Garcia-Vila, M., Gayler, S., Hoogenboom, G., Hunt, L.A., Izaurralde, R.C., Jabloun, M., Jones, C.D., Kersebaum, K.C., Kimball, B. A., Koehler, A.-K., Kumar, S.N., Nendel, C., O'Leary, G.J., Olesen, J.E., Ottman, M.J., Palosuo, T., Prasad, P.V.V., Priesack, E., Pugh, T.A.M., Reynolds, M., Rezaei, E.E., Rötter, R.P., Schmid, E., Semenov, M.A., Shcherbak, I., Stehfest, E., Stöckle, C.O., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Thorburn, P., Waha, K., Wall, G.W., Wang, E., White, J.W., Wolf, J., Zhao, Z., Zhu, Y., 2016. Similar estimates of temperature impacts on global wheat yield by three independent methods. Nat. Clim. Chang. 6, 1130–1136.

Lobell, D.B., 2007. Changes in diurnal temperature range and national cereal yields. Agric. For. Meteorol. 145 (3-4), 229–238.

Lobell, D.B., Schlenker, W., Costa-Roberts, J., 2011. Climate trends and global crop production since 1980. Science 333 (6042), 616–620.

Long, S.P., Ainsworth, E.A., Leakey, A.D.B., Nösberger, J., Ort, D.R., 2006. Food for thought: lower-than-expected crop yield stimulation with rising CO2 concentrations. Science 312 (5782), 1918–1921.

Luterbacher, J., Dietrich, D., Xoplaki, E., Grosjean, M., Wanner, H., 2004. European seasonal and annual temperature variability, trends, and extremes since 1500. Science 303 (5663), 1499–1503.

Marra, M.C., Schurle, B.W., 1994. Kansas wheat yield risk measures and aggregation: a meta-analysis approach. J. Agr. Resource Econ. 19 (1), 69–77.

Meier, U. (Ed.), 2001. Growth Stages of Mono- and Dicotyledonous Plants, 2nd ed. Berlin and Braunschweig.

Mendelsohn, R., Nordhaus, W.D., Shaw, D., 1994. The impact of global warming on agriculture: a Ricardian analysis. Am. Econ. Rev. 84 (4), 753–771.

Miao, R., Khanna, M., Huang, H., 2016. Responsiveness of crop yield and acreage to prices and climate. Am. J. Agr. Econ. 98 (1), 191–211.

Millo, G., 2014. Robust standard error estimators for panel models: a unifying approach. In: MPRA Working Paper. No. 54954. University Library of Munich, Germany. https://ideas.repec.org/p/pra/mprapa/54954.html.

Müller, C., Robertson, R.D., 2014. Projecting future crop productivity for global economic modeling. Agr. Econ. 45 (1), 37–50.

Odening, M., Shen, Z., 2014. Challenges of insuring weather risk in agriculture. Agr.

Finance Rev. 74 (2), 188–199.
Osborne, T.M., Wheeler, T.R., 2013. Evidence for a climate signal in trends of global crop yield variability over the past 50 years. Environ. Res. Lett. 8 (2), 1–9.
Oury, B., 1965. Allowing for weather in crop production model building. J. Farm Econ. 47 (2), 270–283.
Pardey, P.G., Chan-Kang, C., Dehmer, S.P., Beddow, J.M., 2016. Agricultural R&D is on the move. Nature 537 (7620), 301–303. 15 September.
Perkins, S.E., 2015. A review on the scientific understanding of heatwaves - their measurement, driving mechanisms, and changes at the global scale. Atmos. Res., 242–267.
Ray, D.K., Gerber, J.S., MacDonald, G.K., West, P.C., 2015. Climate variation explains a third of global crop yield variability. Nat. Commun. 6 (5989), 1–9.
Reidsma, P., Ewert, F., Lansink, A.O., 2007. Analysis of farm performance in Europe under different climatic and management conditions to improve understanding of adaptive capacity. Clim. Chang. 84 (3–4), 403–422.
Roberts, M.J., Schlenker, W., Eyer, J., 2013. Agronomic weather measures in econometric models of crop yield with implications for climate change. Am. J. Agr. Econ. 95 (2), 236–243.
Rötter, R., van de Geijn, S.C., 1999. Climate change effects on plant growth, crop yield and livestock. Clim. Chang. 43 (4), 651–681.
Schär, C., Vidale, P.L., Lüthi, D., Frei, C., Häberli, C., Liniger, M.A., Appenzeller, C., 2004. The role of increasing temperature variability in European summer heatwaves. Nature 427, 332–336.
Schlenker, W., Lobell, D.B., 2010. Robust negative impacts of climate change on African agriculture. Environ. Res. Lett. 5 (1), 1–8.
Schlenker, W., Roberts, M.J., 2009. Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. Proc. Natl. Acad. Sci. 106 (37), 15594–15598.
Schulze Steinmann, M., Holm-Müller, K., 2010. Thünensche Ringe der Biogaserzeugung – der Einfluss der Transportwürdigkeit nachwachsender Rohstoffe auf die Rohstoffwahl von Biogasanlagen. Germ. J. Agr. Econ. 59 (1), 1–12.
Siebert, S., Ewert, F., 2012. Spatio-temporal patterns of phenological development in Germany in relation to temperature and day length. Agric. For. Meteorol. 152, 44–57.
Statistisches Bundesamt, 2014. Index der Einkaufspreise landwirtschaftl. Betriebsmittel: Deutschland, Wirtschaftsjahr, Messzahlen mit/ohne Umsatzsteuer, Landwirtschaftliche Betriebsmittel. Tabellencode: 61221-0002.
https://www-genesis.destatis.de/genesis/online (Accessed 29.10.2014).
Statistisches Bundesamt, 2015. Anbaufläche (Feldfrüchte und Grünland): Deutschland, Jahre, Fruchtarten. Tabellencode 41241-0001. https://www- genesis.destatis.de (Accessed 23.07.2015).
Surminski, S., Bouwer, L., Linnerooth-Bayer, J., 2016. How insurance can support climate resilience. Nat. Clim. Chang. 6, 333–334.
Tannura, M.A., Irwin, S.H., Good, D.L., 2008. Weather, Technology and Corn and Soybean Yields in the U.S. Corn Belt. Marketing and Outlook Research Report 2008-01. Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign.
Ward, P.S., Florax, R.J.G.M., Flores-Lagunes, A., 2014. Climate change and agricultural productivity in Sub-Saharan Africa: a spatial sample selection model. Europ. Rev. Agr. Econ. 41 (2), 199–226.
Wessolek, G., Asseng, S., 2006. Trade-off between wheat yield and drainage under current and climate change conditions in northeast Germany. Eur. J. Agron. 24 (4), 333–342.
Wheeler, T., von Braun, J., 2013. Climate change impacts on global food security. Science 341 (6145), 508–513.
Woodard, J.D., Garcia, P., 2008. Weather derivatives, spatial aggregation, and systemic risk: implications for reinsurance hedging. J. Agr. Resource Econ. 33 (1), 34–51.
Wooldridge, J.M., 2009. Introductory Econometrics. A Modern Approach. South-Western Cengage Learning, Mason, Ohio.
Xu, W., Filler, G., Odening, M., Okhrin, O., 2010. On the systemic nature of weather risk. Agr. Finance Rev. 70 (2), 267–284.
Yang, S.-R., Koo, W.W., Wilson, W.W., 1992. Heteroskedasticity in crop yield models. J. Agr. Resource Econ. 17 (1), 103–109.
You, L., Rosegrant, M.W., Wood, S., Sun, D., 2009. Impact of growing season temperature on wheat productivity in China. Agric. For. Meteorol. 149, 1009– 1014.

3 Level normalized modeling approach of yield volatility for winter wheat and silage maize on different scales within Germany

Niveauneutrale Modellierung der Ertragsvolatilität von Winterweizen und Silomais auf mehreren räumlichen Ebenen in Deutschland

Christoph Gornott[1*] und Frank Wechsung[1]

[1] Potsdam Institute for Climate Impact Research (PIK)

[*] Corresponding author

3.1 Abstract

Weather-related yield volatility is an important production risk for agriculture. Especially, negative yield anomalies could increase through climate change. We develop and investigate statistical crop yield models which can be used to predict crop yield impacts of weather and climate projections. The models are applied to winter wheat and silage maize, which are the most important annual crops as winter and spring crops, respectively, in Germany. The yields of both crops were modeled on county level, but evaluated on federal state, river basin or national level. We use three regression methods: separate time series model, panel data model, and random coefficient model. Within the Cobb-Douglas production function and first difference transformation, changing rates (of yield and factor anomalies) are related to each other. To include the conditions of vegetative and generative plant development, we use weather variables summed to quarter- and half-year values. Furthermore, our models are controlled with proxy variables for economic impacts to estimate unbiased weather parameters. Our study shows that the simple separate time series models explain (measured by the Nash-Sutcliffe model efficiency coefficient) yield anomalies best. They perform generally better (0.81) than the panel data models (0.72) due to a more accurate reproduction of exceptional yield changes at the county level. The random coefficient models performed between the separate time series models and panel data models (0.78). The aggregation of county yields to federal statement and river basin yields improves the model accuracy by +0.14. The aggregation effect is at highest for the panel data model on river basin scale (+0.26). The models for both crops achieve a similar goodness of fit. The spatial distribution of model parameters reflects the prevailing soil and weather characteristics within Germany relevant for the different plant development periods. Our statistical models capture collinear factors within yield formation. These are, for example, pests and diseases, or the adaptation behavior of farmers on changing climatic or economic conditions. Due to the normalization, the yield changes are independent of technological levels and can be combined with weather and climate projection without any bias correction. The coarse temporal subdivision of the weather variables supports robust assessments of climate change projections. To conclude, our models are suitable for the combination of yield assessments with weather and climate projections, because they reproduce yields from out-of-sample years robustly. In general, the separate time series models reproduce best the measured yield changes.

Keywords: statistical crop yield models, climate impacts, winter wheat, silage maize, volatility

3.2 Zusammenfassung

Wetterbedingte Ertragsschwankungen stellen für die Landwirtschaft ein Produktionsrisiko dar. Besonders problematisch sind dabei negative Ertragsanomalien, die sich durch den Klimawandel häufen können. Im Rahmen dieser Studie wurden statistische Ertragsmodelle entwickelt und getestet, mit denen Ertragsanomalien modelliert und fortgeschrieben werden können. Für die Modellierung wurden als winterannuelle Kultur Winterweizen und als sommerannuelle Kultur Silomais als die Kulturen mit dem jeweils größten Anbauumfang in Deutschland ausgewählt. Die Erträge dieser beiden Kulturen wurden auf Landkreisebene modelliert und dann auf der Ebene der Bundesländer, Flusseinzugsgebiete und für Deutschland verglichen. Dazu wurden drei statistische Ansätze verwendet: separate Zeitreihenmodelle, Paneldatenmodelle und Zufallskoeffizientenmodelle. Über die funktionale Form der Cobb-Douglas-Produktionsfunktion und die Variablentransformation wurden Änderungen im Vergleich zum Vorjahr (Ertrags- und Faktoranomalien) miteinander in Beziehung gesetzt. Halb- und vierteljährlich summierte Witterungsvariablen gingen in die Modellbildung ein. Den Witterungseinfluss verzerrende ökonomische Einflüsse wurden von Proxyvariablen quantifiziert. Die Ergebnisse (gemessen am Nash-Sutcliffe Modell-Effizienz-Koeffizienten) der Studie zeigten, dass die methodisch einfachsten separaten Zeitreihenmodelle Ertragsanomalien durchgehend besser (0.81) erklärten als die Paneldatenmodelle (0.72) und auch außergewöhnliche, landkreisindividuelle Ertragsänderungen erfassten. Die Erklärungskraft der Zufallskoeffizientenmodelle lag zwischen den separaten Zeitreihenmodellen und den Paneldatenmodellen (0.78). Durch die Aggregation der Landkreiserträge zu Flusseinzugsgebiets- und Bundesländererträgen wurden höhere Erklärungswerte erreicht (+0.14). Dieser Aggregationseffekt war am höchsten beim Paneldatenmodell für Flusseinzugsgebiete (+0.26). Für beide Kulturen werden ähnliche Erklärungswerte erreicht. Die räumliche Verteilung der Modellparameter spiegelte die vorherrschenden Boden- und Klimaeigenschaften Deutschlands in den unterschiedlichen Entwicklungsperioden wieder. Durch die Normierung sind die Erträge einerseits unabhängig vom technologischen Niveau, andererseits können sie ohne Fehlerkorrektur direkt mit simulierten Wetter- und Klimamodellen kombiniert werden. Durch die grobe zeitliche Einteilung der Witterungsvariablen lassen sich mit den Modellen robuste Projektionen abgeben. Unsere statistischen Modelle

erfassten kollinear verlaufende Faktoren der Ertragsbildung, beispielsweise Schädlinge oder das Anpassungsverhalten der Landwirte an sich ändernde klimatische oder ökonomische Bedingungen. Dadurch konnten sie Praxiserträge besser abbilden als prozessbasierte Modelle. Die geschätzten statistischen Modelle sind geeignet, um Ertragsanomalien für Wetter- und Klimaprojektionen fortzuschreiben. Die separaten Zeitreihenmodelle reproduzierten insgesamt am besten die gemessenen Ertragsänderungen.

Stichwörter: Statistische Ertragsmodelle, Klimafolgen, Winterweizen, Silomais, Volatilität

3.3 Einleitung

In den Anfangsjahren der Ertragsmodellierung waren statistische Modelle zunächst die einzige praktikable Möglichkeit, witterungsbedingte Schwankungen historischer landwirtschaftlicher Erträge zu modellieren (Doll, 1967; Oury, 1965; Shaw, 1964). Statistische Modelle erklären die Volatilität der endogenen Variable Ertrag aus der Volatilität ertragsrelevanter exogener Variablen (z.B. Niederschlag). Innerhalb einer funktionalen Form werden die ertragserklärenden Parameter der Modelle geschätzt. Die exogenen Variablen sind häufig auf Witterungseinflüsse begrenzt (Iizumi et al., 2013; Lobell und Asner, 2003; Roberts et al., 2012; Schlenker und Roberts, 2009). Sie können aber neben den witterungsbedingten auch pedosphärische und ökonomische (Adaption, Faktor- und Produktpreise) Ertragseinflüsse berücksichtigen (Bakker et al., 2005; Reidsma et al., 2007; You et al., 2009).

Mit der Entwicklung moderner Rechentechnik gewannen prozessbasierte Ertragsmodelle an Bedeutung (Nendel et al., 2013; Tannura et al., 2008). Diese Modelle zerlegen die pflanzenphysiologische Ertragsbildung in Teilprozesse. Theoretisch können prozessbasierte Modelle Ertragsvolatilität für ein weites Spektrum von Umweltbedingungen, mit nur geringen Änderungen im Parametersatz, simulieren (Asseng et al., 2013; Palosuo et al., 2011). Bei den Parametern dieser Modelle handelt es sich nicht, wie bei statistischen Modellen, um Schätzungen, sondern um experimentell begründete Setzungen. Räumliche unterschiedliche Erträge werden bei prozessbasierten Modellen also ausschließlich über die Variation der Variablen, aber nicht der Parameter errechnet.

In den letzten Jahren haben statistische Modelle wieder an Bedeutung gewonnen, da sie, im Gegensatz zu prozessbasierten Modellen, schwer modellierbare Faktoren implizit berücksichtigen können. Zu diesen Faktoren gehören u.a. die Wirkungen von Krankheiten und Schädlingen, vorhandener Agrotechnik oder Faktorpreisen, welche (witterungsabhängig und witterungsunabhängig) die Bestandsführung und die Anpassungsreaktion der Landwirte an den Klimawandel beeinflussen (Lobell und Burke, 2010). Diese von prozessbasierten Modellen nur mit hohem Aufwand zu erfassenden Faktoren können die direkte Klimawirkung auf den Ertrag überlagern (Challinor et al., 2014). Dies kann Fehleinschätzungen begünstigen.

Ziel der vorliegenden Studie war es, statistische Ertragsmodelle für die Abschätzung von Witterungseinflüssen auf die Erträge landwirtschaftlicher Kulturen zu entwickeln und zu testen. Dabei standen die in Deutschland angebauten Kulturen Winterweizen (*Triticum aestivum L.*) und Silomais (*Zea mays L.*) im Mittelpunkt. Winterweizen ist die bedeutendste winterannuelle, Silomais die bedeutendste sommerannuelle Kultur Deutschlands (Statistisches Bundesamt, 2013). Zudem repräsentiert der Winterweizen stellvertretend den Witterungseinfluss auf C_3 Pflanzen, mit dem Silomais lässt sich der Witterungseinfluss auf C_4 Pflanzen ableiten. Die Ergebnisse zeigen und diskutieren wir auf Modellierungsebene der Landkreise und räumlich aggregiert. Aggregationsebenen sind Deutschland, seine Bundesländer und Flusseinzugsgebiete. Sie werden nachfolgend als Nation und Subnation bezeichnet. Zusammen werden Nation (d.h. Deutschland) und Subnationen (d.h. Bundesländer und Flusseinzugsgebiete) als (Sub)Nation(en) bezeichnet.

Die räumliche Heterogenität der Witterungs-Ertragseinflüsse innerhalb der (Sub)Nation(en) lässt sich statistisch über verschiedene Methoden der Parameterschätzung berücksichtigen: indirekt, mit räumlich separat geschätzten Zeitreihenmodellen (STSM), oder direkt, mit Paneldatenmodellen (PDM) oder Zufallskoeffizientenmodellen (RCM). Mit den methodisch sehr simplen STSMs werden die Parameter der Landkreismodelle separat und unabhängig voneinander geschätzt (Lobell und Burke, 2010). Durch die separate Schätzung der STSMs wird die räumliche Heterogenität innerhalb der (Sub)Nation(en) erfasst. PDMs hingegen schätzen zeitliche und räumliche Ertragsänderungen direkt über einen (für alle Landkreise geltenden) Parametersatz je (Sub)Nation (You et al., 2009). Individuelle Witterungs-Ertragseinflüsse können von den PDM-Parametern daher nur (sub)national (Bundesländer, Flusseinzugsgebiete, Deutschland) und nicht auf Landkreisebene erfasst werden. RCMs belegen eine Zwischenposition bei der Parameterschätzung. Sie schätzen auf (sub)nationaler Ebene mit landkreisindividuellen Parametern (Reidsma et al., 2007). Dafür benötigen sie allerdings ein methodisch komplexeres Schätzverfahren.

Durch die Trennung von Modellierungsebene (Landkreise) und Betrachtungsebene (Bundesländer, Flusseinzugsgebiete, Deutschland) wird gezielt die Wirkung von Aggregation genutzt (Woodard und Garcia, 2008). Durch Aggregation werden Einflüsse (z.B. Krankheiten und Schädlinge) herausgefiltert, die auf Landkreisebene die Modelle verzerren. Durch das Filtern kann sich auf den aggregierten Ebenen eine höhere Modellgüte als auf der eigentlichen Modellierungsebene ergeben. Dennoch können über diesen Ansatz landkreisindividuelle Witterungseinflüsse auf den Ertrag erfasst werden (Butler und Huybers, 2013).

Ein Vorzug statistischer Modelle besteht darin, dass sie sich im Unterschied zu prozessbasierten Modellen von den zugrundeliegenden absoluten Niveaus lösen können. Es werden dann nicht mehr die absoluten Erträge, sondern absolute oder relative Änderungen gegenüber dem Vorjahr (erste Differenzen, Quotienten) errechnet. Dies hat sowohl für die Schätzung, als auch für die Anwendung der Modelle Vorteile und Nachteile. Im Vorfeld der Schätzung können durch die Bildung von Differenzen oder Quotienten Ertragstrendeinflüsse oder Änderungen des Trends, durch züchterischen und technologischen Fortschritt, eliminiert werden (Lobell und Asner, 2003; You et al., 2009). Beispielsweise ist durch die Niveaunormierung die Stagnation des Ertragstrends in Deutschland, wie von Brisson et al. (2010) gezeigt, unproblematisch für die Schätzung in statistischen Modellen. Aber auch systematische Fehler der Variablen verlieren an Bedeutung. Eine explizite Fehlermodellierung bei der Verwendung von Daten aus Klimamodellen wird überflüssig. Selbst wenn die Klimamodelle einen systematischen Fehler aufweisen, kann der Ertragseinfluss von Klimaänderungen abgeschätzt werden (Lobell, 2013). Nachteilhaft ist, dass durch die Trennung vom absoluten Niveau nur noch begrenzt Aussagen über dessen Veränderung gemacht werden können (da dieses nicht mehr in den Daten und Parametern enthalten ist). Zudem wird bei der Schätzung nicht mehr die Niveauabhängigkeit in den Relationen zwischen Ertrag und Ertragsfaktoren berücksichtigt.

Für die von uns geprüften Modellansätze verwendeten wir, analog zu Oury (1965), die Cobb-Douglas-Produktionsfunktion als funktionale Form. In dieser werden logarithmierte erste Differenzen zueinander in Beziehung gesetzt. Die logarithmierten ersten Differenzen der Erträge bzw. der ertragsbeeinflussenden Faktoren werden nachfolgend vereinfachend als relative Ertrags- bzw. Faktoränderung bezeichnet. Die Cobb-Douglas-Produktionsfunktion berücksichtigt Substitution und Interaktion zwischen den exogenen Variablen. Zudem sind die Parameter als relative Ertragsänderungen zu interpretieren und daher direkt miteinander vergleichbar (Wooldridge, 2013: 351-354). Eine Vielzahl anderer funktionaler Formen ist möglich. Wir beschränken uns auf die Cobb-Douglas-Produktionsfunktion, da sie sich sowohl in ökonomischen (You et al., 2009) als auch pflanzenbaulichen Anwendungen (Lee et al., 2013) bewährt hat.

Da das Wetter in den phänologischen Entwicklungsphasen unterschiedlich auf den Ertrag wirkt (Chmielewski et al., 2004), ist eine zeitliche Aufteilung der Wachstumsperiode sinnvoll. Dixon et al. (1994) und Lobell et al. (2011) teilen die Witterungsvariablen phänologisch nach Kalendermonaten ein. Chmielewski und Köhn (2000) verwenden für ihre Ertragskomponentenanalyse von Winterroggen fünf phänologische Entwicklungsperioden unabhängig von den Kalendermonaten. Moore und Lobell (2014), Butler und Huybers (2013) und You et al. (2009) unterteilen die Wachstumsperiode nicht. Wir nutzten eine vergleichsweise grobe Unterteilung nach Viertel- und Halbjahren, um partielle Witterungswirkungen während der vegetativen und generativen Entwicklung abzubilden. Diese grobe Unterteilung erachteten wir als ausreichend für die Anwendung der Modelle zur Abschätzung von Klimafolgen, da Klimasimulationen robuste Tendenzen erst bei gröberer zeitlicher Aggregation erkennen lassen.

Die Einteilung der Witterungsvariablen basiert auf zwei Haupteinflussfaktoren. Der erste Einflussfaktor ist die durch die Globalstrahlung (R_s) ankommende Energie. Diese bestimmt das potenzielle Wachstum der Pflanze (Monteith, 1977). Faktoren, die dieses potenzielle Wachstum negativ beeinflussen lassen sich als Stressfaktoren (zweiter Einflussfaktor) beschreiben. Besonders sensitiv reagieren Pflanzen in Deutschland auf eine unzureichende Wasserversorgung (Chmielewski und Köhn, 2000; Kersebaum und Nendel, 2014). Die Wasserversorgung wurde durch die Variablen Niederschlag und potenzieller Evapotranspiration (*ETP*) abgebildet. Andere Stressfaktoren, wie z.B. hohe Temperaturen (Lobell et al., 2013) werden nicht berücksichtig, da sie vergleichsweise selten in Deutschland wirksam werden.

Zwischen den Witterungsvariablen in einem statistischen Modell kann Multikollinearität, das heißt, eine Korrelation der exogenen Variablen untereinander, auftreten. Beispielsweise ist die R_s hoch mit der Temperatur (Bristow und Campbell, 1984) und aus der Temperatur (T) errechneten Variablen korreliert (Lobell, 2010). Dixon et al. (1994) und Lobell und Asner (2003) verwenden die R_s als ertragserklärende Witterungsvariable. Dixon et al. (1994) zeigen, dass das Weglassen der R_s zwar nur zu einem geringen Verlust an Erklärungskraft führt. Allerdings ändern sich die Parameter beträchtlich (*omitted-variable bias*) und die Modelle verlieren deutlich an Voraussagefähigkeit (gemessen am RMSE). Als Proxyvariable für R_s verwenden You et al. (2009) den Bewölkungsgrad. Wir verwendeten eine temperaturnormierte Globalstrahlung (*SRT*), um einerseits den Strahlungseinfluss auf den Ertrag abzubilden und andererseits Kollinearität mit temperaturabhängigen Variablen, wie Sättigungsdefizit (*VPD*) oder *ETP* zu mindern.

Kaufmann und Snell (1997) diskutieren, dass nicht im Modell berücksichtigte, aber ertragsrelevante (ökonomische) Variablen zu einem *omitted-variable bias* führen. Ertragseffekte durch sich ändernde ökonomische Bedingungen können die Ertragswirkungen der interannuellen Witterungsänderungen überlagern (Reidsma et al., 2007). Gerade auf weniger produktiven Ertragsstandorten haben sich die Anbaufläche und die Intensität des Faktoreinsatzes durch gesetzliche Flächenstilllegungsquoten und über die Förderung von Biogas und Biodiesel geändert (Krause, 2008). Aus diesem Grund sind konstante Bewirtschaftungsbedingungen über den betrachtenden Zeitraum und innerhalb von Deutschland nicht realistisch (vgl. Lobell und Asner, 2003). Um die ökonomischen Ertragswirkungen abzubilden, nutzten wir als ökonomische Proxyvariablen Düngerpreis und Anbaufläche der jeweiligen Kultur. Der Düngerpreis bildet den Produktionsfaktoreinsatz bei sich ändernder Rentabilität ab. Die im Zeitraum 1991 bis 2010 gestiegene Anbaufläche von Winterweizen (+34%) und Silomais (+40%) steht für die Änderung der Gemeinsamen Agrarpolitik (GAP).

Wir prüften die Leistungsfähigkeit des oben skizzierten Konzeptes zur niveauneutralen Ertragsmodellierung auf eine Anschlussnutzung in der Klimafolgenforschung. Es wurde gezielt die Wirkung von Aggregation genutzt. Die Parameterschätzung wurde auf Landkreisebene vorgenommen, die Modellevaluierung erfolgte auf (sub)nationaler Ebene (Bundesland, Flussgebiet und Deutschland). Die Vor- und Nachteile von drei statistischen Ansätzen (STSMs, PDMs, RCMs) wurden dargestellt und diskutiert.

3.4 Material und Methoden

3.4.1 Datengrundlage und Aggregation der Variablen

Die Ertragsdaten wurden landwirtschaftlichen Ertragsstatistiken der deutschen Landkreise von 1991 bis 2010 entnommen. Die witterungsbedingten Einflüsse auf den Winterweizen- und Silomaisertrag wurden über die Witterungsvariablen *ETP*, *SRT* und Niederschlag (*Nied*) abgebildet (Tab. 1). Die Witterungsvariablen wurden aus den Daten der einzelnen, über Deutschland verteilten, Wetterstationen berechnet. Aus den täglichen Stationsdaten wurden zunächst die genannten Variablen berechnet, dann wurden diese zu Halb- und Vierteljahreswerten summiert und schließlich für jeden Landkreis gemittelt. Die ökonomischen Proxyvariablen Anbaufläche der jeweiligen Kultur und Düngerpreis liegen nur für Gesamtdeutschland vor. Sie flossen deshalb nicht landkreisindividuell in die Modelle ein. Eine ausführlichere Beschreibung der Datenbasis befindet sich im Appendix A.1.

Tab. 1. Genutzte exogene Variablen für das Winterweizen- (WW) und das Silomaismodell (SM). Die Variablen unterscheiden sich hinsichtlich ihrer zeitlichen Einteilung (Periode), der Aggregation (Agg) und nach der räumliche Beobachtungsebene (Ebene mit LK = Landkreis und Nat = National). Die Perioden stehen für die Kalendermonate November bis April (Nov-Apr), Mai bis Juli (Mai-Juli) und August bis Oktober (Aug-Okt). Die abgekürzten exogenen Variablen sind potenzielle Evapotranspiration (ETP), temperaturnormierte Globalstrahlung (SRT). Das Basisjahr des Indexes ist 2005=100.

Variable	Kultur	Periode	Agg	Ebene
ETP	WW	Nov-Apr	Summe	LK
ETP	WW, SM	Mai-Juli	Summe	LK
ETP	SM	Aug-Okt	Summe	LK
Niederschlag	WW	Nov-Apr	Summe	LK
Niederschlag	WW, SM	Mai-Juli	Summe	LK
Niederschlag	SM	Aug-Okt	Summe	LK
SRT	WW, SM	Mai-Juli	Summe	LK
Düngerpreis	WW, SM	Jan-Dez	Index	Nat
Ackerfläche WW	WW	Jan-Dez	Index	Nat
Ackerfläche SM	SM	Jan-Dez	Index	Nat

3.4.2 Berechnungsgrundlagen von Witterungsvariablen

Die tägliche *SRT* wird nach Gleichung 1 berechnet. Um eine Division durch Null zu vermeiden wurde die Temperaturachse um den Summanden 20 korrigiert (vgl. Oury, 1965, Ariditätsindex nach de Martonne).

$$SRT = \frac{R_S}{T_{avg} + 20}, \qquad \text{mit} \tag{1}$$

R_S – tägliche Globalstrahlungssumme [J cm^{-2}],
T_{avg} – durchschnittliche Tagestemperatur [°C],
SRT – temperaturnormierte Globalstrahlungssumme [J $°C^{-1}$ cm^{-2}].

Die *ETP* (Gleichung 2), berechnet nach Haude, setzt sich aus zwei Teilen zusammen: dem Haude-Faktor f_H und dem *VPD* (Bormann et al., 1996; Schrödter, 1985). Da die f_H für Mais und Weizen nicht für die gesamte Wachstumsphase dieser Kulturen verfügbar sind, verwendeten wir für f_H das arithmetische Mittel der f_H-Werte von Weizen, Mais und Grünland. Diese berücksichtigen die spezifischen Eigenschaften von Weizen und Mais und sind monatsweise für das ganze Jahr verfügbar.

$$ETP = f_H\, VPD, \qquad \text{mit} \tag{2}$$

f_H – Haude Faktor,
VPD – tägliches Sättigungsdefizit [hPa].

Das *VPD* (Gleichung 3) errechnet sich nach der Magnus-Formel aus Maximumtemperatur (T_{max}) und Taupunkttemperatur (DVWK, 1996; Sonntag, 1990). Näherungsweise kann statt der Taupunkttemperatur die tägliche Minimumtemperatur (T_{min}) verwendet werden (Castellvi et al., 1996; Roberts et al., 2012). Den von Roberts et al. (2012) und Donatelli et al. (2006) genutzten Skalierungsfaktor ersetzten wir durch den Faktor der Magnus-Formel (6.11) nach Sonntag (1990).

$$VPD = 6.11\left(e^{\left(\frac{17.269\, T_{max}}{237.3\, +\, T_{max}}\right)} - e^{\left(\frac{17.269\, T_{min}}{237.3\, +\, T_{min}}\right)}\right) \tag{3}$$

3.4.3 Zeitliche Einteilung der Witterungsvariablen

Die zeitliche Aggregation unserer Klimavariablen orientierte sich an der phänologischen Entwicklung von Winterweizen und Silomais in Deutschland. Die Aussaat- und Erntezeiten von Winterweizen und Silomais variieren zwischen den Jahren und den Landkreisen. Von 1992-2010 erfolgt die Aussaat von Winterweizen näherungsweise von Anfang bis Mitte Oktober (±4 Tage) (Auflaufen Ende Oktober) und jene von Silomais gegen Ende April (±5 Tage) (Auflaufen Anfang Mai). Die Erntezeit von Winterweizen ist näherungsweise Anfang August (±12 Tage) und von Silomais Ende September (± ein Monat). Innerhalb von Deutschland variiert die Erntezeit von Weizen um ± einen Monat, von Silomais um ±22 Tage (DWD, 2014). Für die Ertragsmodelle verwendeten wir Viertel- und Halbjahressummen der Witterungsvariablen *ETP*, *SRT* und *Nied*. Diese charakterisieren die Witterungsbedingungen während der vegetativen und generativen Wachstumsphasen der betrachteten Kulturen in Deutschland. In unseren Modellen erstreckte sich die vegetative Phase des Winterweizens vom November des Aussaatjahres bis zum April des Erntejahres, bei Silomais ging sie vom Mai bis zum Juli des Erntejahres. Die generative Phase umfasst die Monate Mai bis Juli beim Winterweizen und August bis Oktober beim Silomais (wir nahmen den auf volle Monate gerundeten spätesten Erntezeitpunkt).

3.4.4 Funktionales Grundmodell

Als Grundmodell wurde die Cobb-Douglas Produktionsfunktion verwendet (Gleichung 4). Mit der Funktion werden die relativen Ertragsänderungen (Änderungen im Vergleich zum Vorjahr) durch die relativen Änderungen der exogenen Variablen geschätzt.

$$\frac{y_t}{y_{t-1}} = \beta_0 \prod_{j=1}^{J} \left(\frac{x_{jt}}{x_{jt-1}}\right)^{\beta_j} \prod_{k=1}^{K} \left(\frac{x_{kt}}{x_{kt-1}}\right)^{\beta_k}, \qquad \text{mit } t = 2, \dots, M, j = 1, \dots, J, k = 1, \dots, K, \tag{4}$$

$\frac{y_t}{y_{t-1}}$ – relative Ertragsänderungen im Jahr *t* in Relation zum Jahr $t-1$ (endogene Variable),

$\frac{x_{jt}}{x_{jt-1}}, \frac{x_{kt}}{x_{kt-1}}$ – relative Änderungen witterungsbedingter (*j*) und ökonomischer (*k*) Ertragsfaktoren im Vergleich zum Vorjahr (exogene Variablen),

β_0 – Parameterwert für den mittleren Trendeinfluss,

β_j, β_k – Parameter für den Faktoreinfluss der *j*-ten witterungsbedingten und *k*-ten ökonomischen Variable. Die Parameter beschreiben den relativen Effekt auf den Ertrag (in Prozent) beim Anstieg einer witterungsbedingten oder ökonomischen Variable um ein Prozent.

3.4.5 Regressionsansätze für Landkreiserträge

Das Grundmodell wurde mit drei unterschiedlichen Regressionsmethoden geschätzt. Die Methoden unterschieden sich hinsichtlich ihrer Berücksichtigung von landkreisindividuellen Ertragseinwirkungen. Zur Vergleichbarkeit der Methoden waren funktionale Form und Transformation für alle Modelle einheitlich. Um die Cobb-Douglas Produktionsfunktion zu linearisieren, wurde das Grundmodell (Gleichung 4) logarithmiert. Durch die Transformation der Modellterme wurden diese für Regressionsmodelle nutzbar: $\log\left((x_t/\, x_{t-1})^{\beta}\right) = \beta \log(x_t/\, x_{t-1}) = \beta(\log x_t - \log x_{t-1})$, der transformierte Modellterm $\log x_t - \log x_{t-1}$ wird nachfolgend als $\Delta \log x_t$ bezeichnet. Zudem wurde das Grundmodell um den Fehler *u* erweitert.

Das STSM (Gleichung 5) schätzt relative Ertragsänderungen der einzelnen Landkreise, $i = 1, \ldots, N$, unabhängig voneinander (Dielman, 1983). Durch die separate Schätzung von *N* Modellen wird im Nachhinein die räumliche Heterogenität berücksichtigt. Der Landkreisindex i ist bei der Schätzung der einzelnen Zeitreihenmodelle noch nicht relevant, sondern hat erst bei der Betrachtung aller separaten Modelle einer (Sub)Nation Bedeutung.

$$\Delta \log y_{it} = \beta_{i0} + \sum_{j=1}^{J} \beta_{ij} \Delta \log x_{itj} + \sum_{k=1}^{K} \beta_{ik} \Delta \log x_{tk} + \Delta \log u_{it}, \quad (5)$$

Im Unterschied zu den STSMs schätzen PDMs (Gleichung 6) direkt die Parameter β über alle Landkreise. Die zeitliche und räumliche Volatilität der Variablen wird als unabhängig voneinander angesehen. Im Fehlerterm schätzen die PDMs landkreisindividuelle, zeitinvariante Einflüsse (z.B. Bodenproduktivität) mit dem unbeobachteten individuellen Fehler (Croissant und Millo, 2008; Wooldridge, 2013).

$$\Delta \log y_{it} = \beta_0 + \sum_{j=1}^{J} \beta_j \Delta \log x_{itj} + \sum_{k=1}^{K} \beta_k \Delta \log x_{tk} + \Delta \log u_{it} \quad (6)$$

RCMs (Gleichung 7) können zwischen STSMs und PDMs eingeordnet werden. In RCMs werden mittlere ($\beta_0, \beta_j, \beta_k$) und landkreisindividuelle Parameter (b_{i0}, b_{ij}, b_{ik}) errechnet. Zusammen ergeben beide Parameter ($\beta + b_i$) den landkreisindividuellen Ertragseffekt β_i. Die Parameter der RCMs können nicht, wie STSMs und PDMs, nach der Methode der kleinsten Quadrate (OLS) geschätzt werden, da die Parameter β_j und b_{ij} nicht unabhängig voneinander sind. Die RCM-Parameter werden daher mit der *Restricted Maximum Likelihood* (REML) Methode geschätzt (Reidsma et al., 2007).

$$\Delta \log y_{it} = (\beta_0 + b_{i0}) + \sum_{j=1}^{J} (\beta_j + b_{ij}) \Delta \log x_{itj} + \sum_{k=1}^{K} (\beta_k + b_{ik}) \Delta \log x_{tk} + \Delta \log u_{it} \quad (7)$$

3.4.6 Aggregation der Modellergebnisse

Die mit den verschiedenen Modellen modellierten Landkreiserträge wurden (über das arithmetische Mittel) zu (sub)nationalen Erträgen aggregiert (Gleichung 8) und mit den aggregierten Beobachtungen verglichen.

$$\Delta \log y_{\bar{\iota}t} = N^{-1} \sum_{i=1}^{N} \Delta \log y_{it} \quad (8)$$

3.4.7 Kreuzvalidierung, Modellgüte und statistische Tests

Wir nutzten eine Kreuzvalidierung, um die Vorhersagefähigkeit der Modelle zu testen. Dazu nahmen wir jeweils alle Beobachtungen eines Jahres aus dem Datensatz heraus, schätzten das Modell und simulierten dann den Ertrag für das fehlende Jahr. Die Reproduzierbarkeit des jeweiligen Jahresertrages wurde für den aggregierten Ertrag, d.h. auf der Ebene der (Sub)Nation(en) beurteilt. Diese *out-of-sample* Prognose führten wir sequentiell für alle Jahre durch.

Die Modelle sollen das mittlere Niveau und die Volatilität der beobachteten relativen Ertragsänderungen reproduzieren. Als Gütemaße verwendeten wir dafür den *Root-Mean-Square Error* (RMSE), das Bestimmtheitsmaß (R^2), das korrigierte Bestimmtheitsmaß (Adj. R^2) und den Nash-Sutcliffe Modell-Effizienz-Koeffizienten (NSE). Der RMSE gibt die absolute, mittlere Fehlschätzung an und berücksichtigt daher unterschiedlich langen Zeitreihen. Die Gütemaße R^2 und Adj. R^2 geben Auskunft wie gut die relative Volatilität der relativen Ertragsänderungen reproduziert wird. Sie zeigen keine Fehlschätzungen des Niveaus an. Das R^2 errechnet sich aus den quadrierten Korrelationskoeffizienten zwischen gemessenen und geschätzten relativen Ertragsänderungen für die Ebene der (Sub)Nation(en). Das Adj. R^2 berücksichtigt zusätzlich die Anzahl der in das Modell aufgenommenen Variablen. Zur Vergleichbarkeit von PDMs und STSMs verwendeten wir für die Berechnung des Adj. R^2 die kritischeren Freiheitsgrade der STSMs. Der NSE reagiert sowohl auf Fehlschätzungen des Niveaus als auch der Schwankungen. Daher ist das NSE auch bei nicht OLS-Bedingungen oder für nicht in die Schätzung einfließende relative Ertragsänderungen (Validierungsergebnisse) nutzbar. Die von Krause et al. (2005) angeführte Überschätzung des NSE bei Extremwerten ist durch die logarithmierten Erträge unproblematisch.

Für STSMs und PDMs testeten wir statistisch die Zulässigkeit von OLS-Schätzern. Die nachfolgenden statistischen Tests werden von Croissant und Millo (2008) und Wooldridge (2013) näher beschrieben. Die funktionale Form prüften wir mit dem *Regression Equation Specification Error Test* (RESET). Mit dem *Lagrange-Multiplier-Test* nach Breusch-Pagan (LM) prüften wir die Modelle auf räumliche Heterogenität. Bei Autokorrelation (Breusch-Godfrey-Test) und/ oder Heteroskedastizität (Breusch-Pagan-Test) sind die Signifikanztests (*t*-Test für Parameter) ungültig. Die Standardfehler der Parameter wurden dann durch robuste Standardfehler nach Arellano ersetzt. Normalverteilung der Residuen testeten wir mit dem Shapiro-Wilk-Test.

Zu den RCMs nach REML werden über die genutzten R-Pakete standardmäßig keine statistischen Tests ausgeboten (siehe Appendix A.2 für eine detaillierte Beschreibung der verwendeten Software). Bei der Schätzung der RCM nach *Maximum Likelihood* werden das *Akaike Information Criterion* (AIC) und das *Bayesian Information Criterion* (BIC) als Gütemaße errechnet. AIC und BIC sind vergleichbar mit dem R^2 und dem Adj. R^2 bei OLS-Schätzern. Letztere sind aber bei REML (hierarchischen Daten) als Gütemaße problematisch in der Anwendung, da bei diesem Schätzverfahren nicht zwangsläufig (wie bei OLS) das mittlere Niveau getroffen wird. Der NSE kann hingegen bei REML verwendet werden. Der NSE hat zudem den Vorteil, dass er als relatives Maß direkt interpretierbar ist.

Die vom BIC vorgenommene Diskriminierung bei zunehmender Variablenanzahl ist für unsere Modelle ohne Bedeutung, da alle Modelle die gleichen Variablen nutzen (Reidsma et al., 2007; Wooldridge, 2013).

3.5 Ergebnisse

3.5.1 Landkreisindividuelle Schätzebene

Die mit den STSMs, PDMs und RCMs simulierten relativen Ertragsänderungen der Landkreise konnten die mittlere Volatilität der beobachteten Weizenerträge reproduzieren. Es gab jedoch Unterschiede zwischen den Modellen bei der Simulation von außergewöhnlich großen Ertragsschwankungen. Dies wird exemplarisch für den Landkreis Oder-Spree (Brandenburg) in Abb. 1 verdeutlicht. Von den STSMs wurden überdurchschnittlich große relative Ertragsänderungen am besten reproduziert. PDMs zeigten hingegen deutliche Schwächen bei der Reproduktion außergewöhnlich großer und nur regional auftretender relativer Ertragsänderungen (im restlichen Brandenburg ist diese außergewöhnliche relative Ertragsschwankung nicht aufgetreten). Die RCMs lagen zwischen den STSMs und den PDMs (NSEs für den Landkreis Oder-Spree: STSM: 0.68, PDM: 0.21, RCM: 0.66, für Brandenburg: STSM: 0.84, RCM: 0.81, PDM: 0.70).

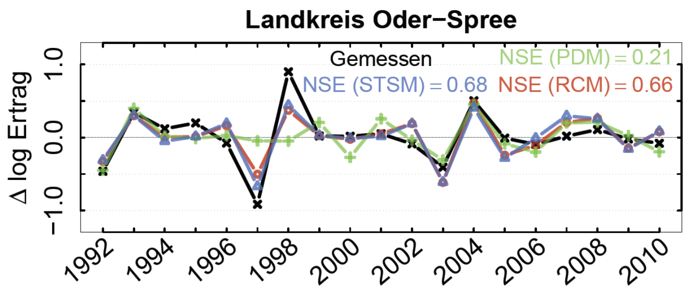


Abb. 1. Gemessene (schwarz) und mit den Modellen STMS, PDM, RCM (blau, grün, rot) geschätzte relative Ertragsänderungen von Winterweizen (Δ log Ertrag) für den Landkreis Oder-Spree (Brandenburg).

3.5.2 Modellgüte nach subnationaler und nationaler Aggregation

Die relativen Ertragsänderungen und Gütemaße für die zu (Sub)Nation(en) aggregierten Modelle sind für STSMs und PDMs exemplarisch in Abb. 2 dargestellt. Die (sub)nationalen relativen Ertragsänderungen der Simulationen und der Kreuzvalidierung wurden nicht systematisch über- oder unterschätzt, da die relativen Ertragsänderungen aus der Kreuzvalidierung sowohl über als auch unter den gemessenen relativen Ertragsänderungen lagen. Die relativen Ertragsänderungen aus der Kreuzvalidierung reagierten nicht sensitiv auf die sich jährlich ändernden Beobachtungen. In den östlichen Regionen erklärten die Simulationen die beobachteten relativen Ertragsänderungen besser als in den westlichen Regionen. Mit steigender Volatilität (hier angegeben als Standardabweichung (SD)) der gemessenen relativen Ertragsänderungen nahm die Erklärungskraft der Modelle je Fruchtart (NSE, R^2, Adj. R^2 und RMSE) in den Simulationen zu. Für beide Kulturen und alle drei Modellansätze galt, dass die Modelle der kontinentalen, nordöstlichen Subnationen (Brandenburg, Mecklenburg-Vorpommern, Elbe, Warnow/ Peene) die höchste Erklärungskraft aufweisen (vgl. auch Tab. 2). Hier besteht bei den gemessenen relativen Ertragsänderungen die höchste Volatilität je Fruchtart. In maritimen (nördlichen) oder bergigen (südlichen) Regionen erklärten die Modelle über die Witterungsvariablen weniger Volatilität als in kontinentalen Regionen. Trotz der deutlichen Unterschiede in der gemessenen Ertragsvolatilität zwischen Winterweizen (WW) und Silomais (SM) in Ostdeutschland (SD: Brandenburg WW: 0.058, SM: 0.082, Mecklenburg-Vorpommern WW: 0.035, SM: 0.056) bestanden zwischen den Kulturen keine Unterschiede in der Erklärungskraft ihrer Modelle (STSM NSE Brandenburg WW: 0.94, SM: 0.94, Mecklenburg-Vorpommern WW: 0.95, SM: 0.93).

Einen vergleichenden Überblick über die erreichte Modellgüte und den Aggregationseffekt geben die Tab. 2 und 3. Diese beziehen sich auf die drei Modelltypen und die beiden Kulturen auf den verschiedenen Aggregationsniveaus. Die STSMs hatten durchgehend höhere NSEs als die PDMs und die NSEs waren auch meistens höher als die der RCMs (Tab. 2). Bei kleinen Flusseinzugsgebieten mit wenigen Landkreisen unterschieden sich die NSEs nicht (z.B. Oder bei Winterweizen). Der winterannuelle Winterweizen erreichte höhere NSEs als der sommerannuelle Silomais auf der subnationalen Ebene, auf der nationalen Ebene waren die NSEs nahezu gleich. Die NSEs aller Modelle waren beim Silomais auf Deutschlandebene (STSM: 0.86, PDM: 0.77, RCM: 0.80) höher als das Mittel der Bundesländer (STSM: 0.75, PDM: 0.68, RCM: 0.73). Beim Winterweizen waren die NSEs der STSM und RCM höher, aber die der PDM geringfügig niedriger. Das Flusseinzugsgebiet Eider ist nur ein Landkreis und daher nicht als PDM oder RCM schätzbar.

Tab. 2. Nash-Sutcliffe Modell-Effizienz-Koeffizienten (NSE) für den Zusammenhang zwischen den gemessenen und simulierten relativen Ertragsänderungen für die Aggregationsebene Bundesländer, Flusseinzugsgebiete und Deutschland. Die Spalten unterscheiden nach Kulturen (Winterweizen (WW) und Silomais (SM)) und Modellen (STMS, PDM und RCM).

	Winterweizen			Silomais		
(Sub)nation	STSM	PDM	RCM	STSM	PDM	RCM
Bundesländer (BL)						
Schleswig-Holstein	0.75	0.67	0.71	0.59	0.51	0.54
Niedersachsen	0.89	0.82	0.86	0.78	0.65	0.72
Nordrhein-Westfalen	0.84	0.8	0.81	0.64	0.55	0.72
Hessen	0.64	0.59	0.6	0.69	0.6	0.65
Rheinland-Pfalz	0.7	0.64	0.68	0.62	0.57	0.6
Baden-Württemberg	0.77	0.66	0.70	0.64	0.54	0.56
Bayern	0.74	0.56	0.65	0.67	0.51	0.57
Saarland	0.70	0.68	0.69	0.77	0.76	0.77
Brandenburg	0.94	0.92	0.93	0.94	0.93	0.94
Mecklenburg-Vorpommern	0.95	0.91	0.93	0.93	0.9	0.92
Sachsen	0.85	0.72	0.79	0.73	0.58	0.72
Sachsen-Anhalt	0.9	0.87	0.89	0.93	0.92	0.93
Thüringen	0.66	0.61	0.63	0.81	0.78	0.79
Flusseinzugsgebiete (FEG)						
Eider	0.78	–	–	0.24	–	–
Schlei/ Trave	0.74	0.65	0.72	0.56	0.52	0.53
Elbe	0.92	0.81	0.88	0.91	0.86	0.88
Weser	0.87	0.82	0.84	0.84	0.77	0.82
Ems	0.85	0.78	0.83	0.58	0.33	0.45
Rhein	0.78	0.69	0.74	0.77	0.68	0.73
Maas	0.70	0.67	0.68	0.76	0.74	0.75
Donau	0.72	0.63	0.66	0.5	0.35	0.39
Warnow/ Peene	0.94	0.91	0.93	0.93	0.89	0.91
Oder	0.76	0.76	0.76	0.93	0.91	0.92
Subnationale Mittelwerte						
Mittel BL	0.79	0.73	0.76	0.75	0.68	0.73
Mittel FEG	0.81	0.75	0.78	0.70	0.67	0.71
National						
Deutschland	0.86	0.71	0.81	0.86	0.77	0.80

In Tab. 3 wird deutlich, dass der Aggregationseffekt (Δ NSE) bei den PDMs für Flusseinzugsgebiete am größten war (WW: 0.26, SM: 0.27). Bei den STSMs und den RCMs war er annähernd gleich (0.10-0.15). Der Aggregationseffekt war auf nationaler Ebene am größten (0.30), gefolgt von den Flusseinzugsgebieten (0.17). Der geringste Aggregationseffekt ergab sich bei den Bundesländern (0.11).

Tab. 3. Aggregationseffekt der aggregierten (Sub)Nationen gegenüber den Landkreisergebnissen als ΔNSE, mit $NSE = NSE_{(Sub)Nation} - \left(N^{-1} \sum_{i=1}^{N} NSE_{Landkreise\ (i)}\right)$. Die Spalten unterscheiden nach Kulturen (Winterweizen (WW) und Silomais (SM)) und Modellen (STMS, PDM und RCM).

	Winterweizen			Silomais		
(Sub)nation	STSM	PDM	RCM	STSM	PDM	RCM
Bundesländer (BL)						
Schleswig-Holstein	0.08	0.01	0.10	0.10	-0.07	0.16
Niedersachsen	0.13	0.16	0.16	0.15	0.07	0.16
Nordrhein-Westfalen	0.16	0.14	0.25	0.14	-0.03	0.31
Hessen	0.07	-0.07	0.14	0.25	0.02	0.28
Rheinland-Pfalz	0.08	-0.02	0.12	0.12	-0.01	0.17
Baden-Württemberg	0.11	0.00	0.11	0.09	-0.04	0.14
Bayern	0.16	-0.10	0.16	0.14	-0.07	0.11
Saarland	0.07	0.02	0.09	0.05	0.18	0.05
Brandenburg	0.09	0.26	0.12	0.08	0.35	0.08
Mecklenburg-Vorpommern	0.13	0.25	0.16	0.12	0.32	0.14
Sachsen	0.09	0.06	0.05	0.13	0.00	0.07
Sachsen-Anhalt	0.05	0.21	0.06	0.11	0.34	0.08
Thüringen	0.10	-0.05	0.18	0.15	0.20	0.17
Flusseinzugsgebiete (FEG)						
Eider	0.00	–	–	0.00	–	–
Schlei/Trave	0.05	0.08	0.06	0.08	0.17	0.13
Elbe	0.19	0.44	0.18	0.20	0.39	0.14
Weser	0.15	0.29	0.19	0.21	0.44	0.21
Ems	0.12	0.23	0.16	0.06	0.19	0.02
Rhein	0.14	0.37	0.20	0.21	0.50	0.31
Maas	0.03	0.14	0.05	0.19	0.33	0.23
Donau	0.18	0.45	0.18	0.04	0.17	0.08
Warnow/Peene	0.11	0.28	0.11	0.12	0.22	0.11
Oder	0.02	0.05	0.04	0.02	0.06	0.03
Subnationale Mittelwerte						
Mittel BL	0.10	0.07	0.13	0.13	0.10	0.15
Mittel FEG	0.10	0.26	0.13	0.11	0.27	0.14
National						
Deutschland	0.20	0.47	0.18	0.28	0.50	0.17

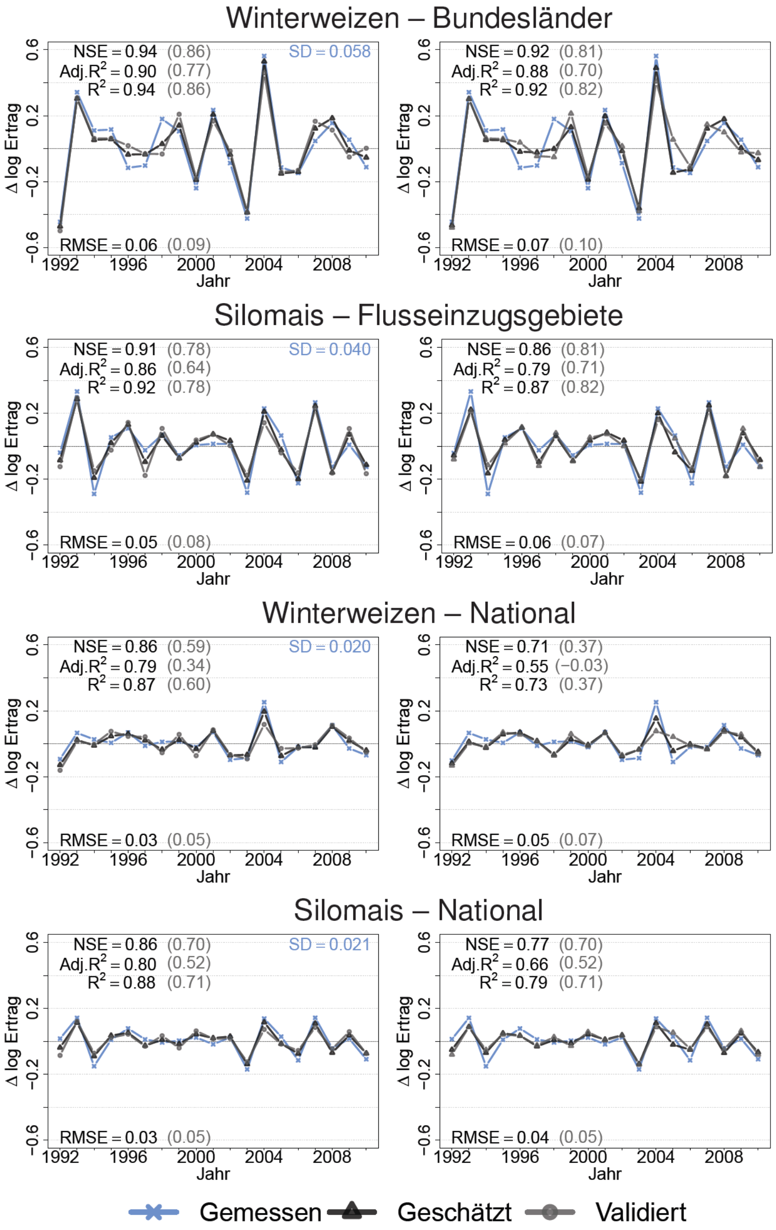


Abb. 2. Gemessene sowie mit zwei Modellen (aggregierte STMS links, PDM rechts) geschätzte und kreuzvalidierte relative Ertragsänderungen (logarithmiert, erste Differenzen) für Winterweizen und Silomais auf verschiedenen regionalen Ebenen. Die relativen Ertragsänderungen sind für Winterweizen auf Bundeslandebene, für Silomais auf Flusseinzugsebene und für beide Kulturen auf nationaler Ebene dargestellt. Die schwarzen Erklärungswerte (NSE, Adj. R², R²) beziehen sich auf die geschätzten Modellergebnisse, die grauen auf die Ergebnisse der Kreuzvalidierung. Die Standardfehler (SD) der gemessenen relativen Ertragsänderungen sind in blau angegeben.

Die räumliche Ausprägung des Aggregationseffektes wird in Abb. 3 am Beispiel der Winterweizen STSMs mit Bezug auf die Korrelation zwischen simulierten und geschätzten relativen Ertragsänderungen gezeigt. Da die relativen Ertragsänderungen nicht systematisch überschätzt waren (vgl. Abb. 2, Tab. 2), ist eine Darstellung, die sich auf den geläufigeren R^2 als Gütemaß bezieht, hinreichend. Subnationen mit geringen R^2s der einzelnen STSMs auf Landkreisebene erreichten nach der Aggregation höhere Erklärungswerte. Unter anderem konnten die relativen Ertragsänderungen der südlichen Subnationen nach der Aggregation besser reproduziert werden. Beispielsweise stieg in der Subnation Bayern das R^2 der STSMs von durchschnittlich 0.58 auf 0.74 nach der Aggregation.

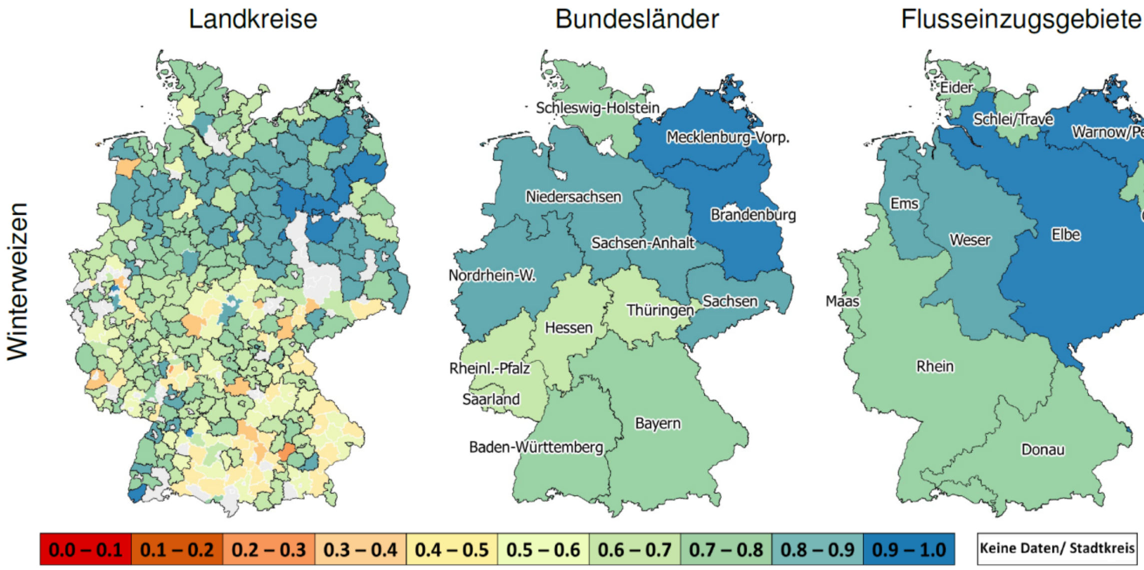


Abb. 3. Räumliche Verteilung derR^2 Werte für den Zusammenhang zwischen gemessenen und mittels STSM geschätzten relativen Ertragsänderungen von Winterweizen auf unterschiedlichen Aggregationsebenen. Die linke Karte zeigt die R^2 der einzelnen STSM auf Landkreisebene (N = 289). Auf der mittleren und der rechten Karte sind die R^2 der aggregierten STSMs für Bundesländer (N = 13) und Flusseinzugsgebiete (N = 10). Die Landkreise mit signifikanten STSM-Modellen (F-Test, $p \leq 0.10$) sind auf der linken Karte schwarz umrandet (N = 198).

3.5.3 Parameter der Witterungsvariablen

Die Parameter der STSMs, PDMs und RCMs sind für die beiden Fruchtarten nach Bundesländern (BL) und Flusseinzugsgebieten (FEG) als *Boxplots* in Abb. 4 dargestellt. Eine große Spannweite der *Boxplots* bedeutet, dass der Ertragseinfluss der Witterung zwischen den Landkreisen (STSM und RCM) bzw. den Subnationen (PDM) räumlich stark variiert. Generell war die Spannweite der Parameter bei den PDMs und den RCMs deutlich kleiner als bei den STSMs.

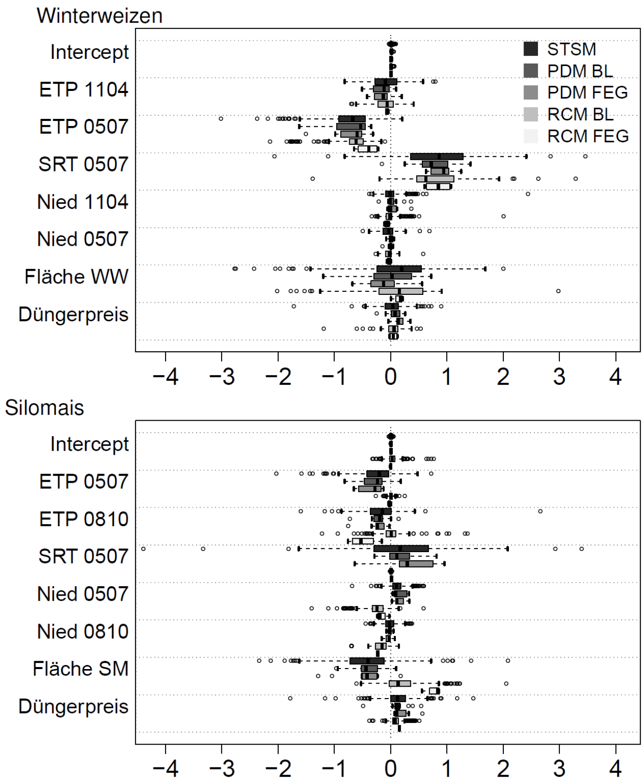


Abb. 4. Geschätzte Koeffizienten der Parameter für alle Modelle für Winterweizen (WW) und Silomais (SM). Die Parameter der STSMs und RCMs beziehen sich auf Landkreise, die Parameter der PDM beziehen sich auf Bundesländer (BL) oder Flusseinzugsgebiete (FEG). Die Zahlen hinter den exogenen Variablen beschreiben die Monate (z.B. 0507 ist Mai bis Juli). Der Balken in der Box ist der Median, die Box repräsentiert das 25% und das 75% Quartil. Die Whiskers sind als Maximum und Minimum definiert, sofern sie kleiner als der 1.5-fache Interquartilsabstand vom Median sind. Ausreißer werden als Punkte dargestellt.

Die räumliche Verteilung für ausgewählte STSM Parameter zeigt exemplarisch Abb. 5. Hier ließen sich klare räumliche Cluster identifizieren. Die Clustergrenzen verliefen intra- und intersubnational, fallen jedoch nur selten mit den Grenzen der Subnationen zusammen. Beim Parameter der *ETP* zeigte sich in den meisten Landkreisen für beide Kulturen ein deutlich negativer relativer Ertragseinfluss (vgl. auch Abb. 4). Am stärksten ($< -1.4\%$) war dieser Einfluss bei Winterweizen von Mai bis Juli im Bundesland Brandenburg und den angrenzenden Landkreisen (relative Ertragsänderung bei einer relativen *ETP* Änderung). Für Silomais ergab sich im Erzgebirge (südliches Sachsen) und im Donaueinzugsgebiet ein neutraler bis leicht positiver Ertragseinfluss bei Zunahme der *ETP*. Niederschlag hatte in der Jugendentwicklung beider Kulturen einen positiven Ertragseinfluss in niederschlagsarmen Regionen, wie Ostdeutschland oder Franken (nördliches Bayern). In Regionen mit mehr Niederschlag hatte dieser einen leicht negativen Einfluss. Das sind in Deutschland der Alpenraum (südliche Grenzen von Deutschland), das Elbsandsteingebirge und die Nordseeküste (nördliche Grenze von Niedersachsen, westliche Grenze von Schleswig Holstein). Insgesamt war der relative Ertragseinfluss der Witterung in kontinentalen Regionen größer, als in maritimen oder (den ackerbaulich genutzten Gebieten) in bergigen Regionen.

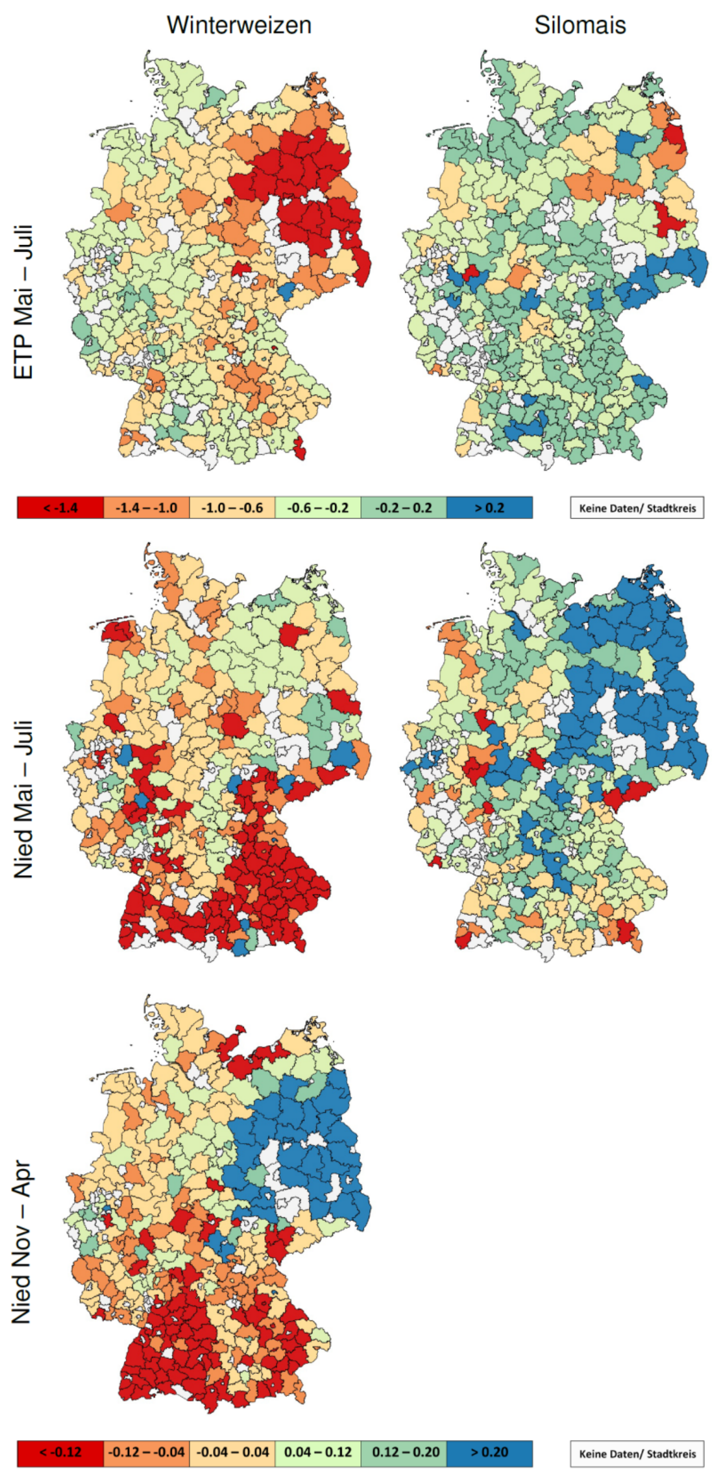


Abb. 5. Räumliche Verteilung der STSM-Parameterkoeffizienten für Winterweizen und Silomais. Die Parameter beschreiben den relativen Ertragseffekt der exogenen Witterungsvariablen *ETP* und *Nied* in den Perioden November bis April und Mai bis Juli.

3.5.4 Multikollinearität der exogenen Variablen und statistische Tests

Die ertragserklärenden Witterungsvariablen zeigten mittelmäßig hohe (0.5-0.7) Korrelationskoeffizienten, dies deutet auf eventuelle Multikollinearität hin. Die Korrelationskoeffizienten waren teilweise statistisch signifikant (das Signifikanzniveau wird über den *p*-Wert beschrieben, dabei ist ***: $p \leq 0.01$, **: $p \leq 0.05$, *: $p \leq 0.1$ und ': $p > 0.1$). Exemplarisch sind nachfolgend Korrelationskoeffizienten der transformierten Witterungsvariablen von Mai bis Juli dargestellt. Die *SRT* war mittelmäßig mit der *ETP* und dem Niederschlag korreliert (0.67^{***}, -0.52^{***}). Die Korrelationen waren statistisch signifikant. Niederschlag und *ETP* waren ebenfalls mittelmäßig, aber signifikant miteinander korreliert (-0.58^{***}). Die Eigenschaften der nicht genutzten R_S sind aber dennoch in der *SRT* enthalten (0.94^{***}). Ein Test auf Multikollinearität der Modelle (*Condition Index*-Test) ergab jedoch, dass es keine Multikollinearität in den Modellen gab. Düngerpreis und Anbaufläche sind geeignete ökonomische Proxyvariablen, da sie den Faktoreinsatz und die Agrarpolitikänderung abbildeten. Die ökonomische Variable Düngerpreis war stark mit anderen Faktorpreisen, wie dem Saatgutpreis (0.86^{***}) und dem Treibstoffpreis (0.83^{***}), korreliert. Der Brotweizenpreis des Vorjahres (Informationsgrundlage des Landwirts für Faktoreinsatz) war in den 90er Jahren sehr gering mit dem Düngerpreis korreliert (-0.14'), nach der GAP Reform 2000 (Liberalisierung der Agrarmärkte) jedoch stark mit diesem korreliert (0.81^{***}). Düngerpreis und Anbaufläche von Winterweizen bzw. Silomais waren miteinander stark korreliert (0.72^{***}/ 0.87^{***}).

Die statistischen Tests geben Aufschluss über die Zulässigkeit der Regressionsmethoden. Für die einzelnen Landkreise, Bundesländer und Flusseinzugsgebiete sind die Ergebnisse der statistischen Tests im Appendix aufgeführt (Abb. S.1). Beim STSM war die genutzte Cobb-Douglas-Funktion als funktionale Form, bis auf wenige Ausnahmen, nicht fehlspezifiziert (RESET). Beim PDM war sie bei Winterweizen teilweise und bei Mais häufig fehlspezifiziert. Weiterhin konnten wir zeigen, dass es in allen Subnationen räumliche Heterogenität gibt (LM-Test) und daher Modelle nötig waren, welche diese berücksichtigen. Autokorrelation und Heteroskedastizität traten bei den STSMs in den überwiegenden Landkreisen nicht auf. Die Modellresiduen waren Normalverteilt. Beim PDM sind aufgrund vorliegender Autokorrelation und/ oder Heteroskedastizität die Parameter häufig ineffizient. Daher verwendeten wir die robusten Standardfehler nach Arellano. Die Parameter mit den robusten Standardfehlern waren überwiegend signifikant. Die NSEs der RCMs werden zusammen mit den NSEs der STSMs und PDMs in Tab. 2 gezeigt.

3.6 Diskussion

3.6.1 Prüffrage und prozessbasierte Modelle

Wir zeigten, dass eine niveauneutrale Ertragsmodellierung der klimabedingten Ertragsvolatilität für die beiden wichtigsten landwirtschaftlichen Kulturen Deutschlands möglich ist. Mit allen drei statistischen Modellansätzen ließ sich die zeitliche und räumliche Volatilität der relativen Ertragsänderungen zufriedenstellend reproduzieren und projizieren. Für die Modellierung waren nur wenige, zeitlich grob aufgelöste Klimavariablen notwendig. Die Ergebnisse verdeutlichten jedoch, dass eine zufriedenstellende Modellqualität nicht auf der untersten Modellierungsebene, sondern erst nach einer räumlichen Aggregation auf Bundesländern, Flusseinzugsgebieten bzw. auf nationale Ebene für Deutschland erreicht wurde. Im Unterschied zu prozessbasierten Modellen waren sowohl ertragsbedingte Entscheidungen der Landwirte (Sortenwahl, Beregnung, Wechsel auf marginale Standorte) als auch die Einflüsse von schwer quantifizierbaren Faktoren (Schädlings- und Krankheitsbefall) implizit berücksichtigt. Diese Eigenschaft stellen auch Lobell und Burke (2010) heraus, jedoch berücksichtigen sie keine ökonomischen Variablen wie beispielsweise You et al. (2009), um kollinear verlaufende Effekte aufzufangen.

Die Modelle können unmittelbar genutzt werden, um Ertragseffekte kurz und mittelfristiger Klimasimulationen abzuschätzen. Hierzu sind allerdings im Grundansatz die Parameter $\beta_0$und β_k auf Null zu setzen. In diesen Parametern spiegeln sich der Trendeinfluss und der Einfluss ökonomischer Proxyvariablen wieder. Im Modellansatz wurden sie eingeführt, um den Einfluss der Witterungsvariablen unverzerrt zu schätzen. Bei den Projektionen spielen sie keine Rolle. Die Projektionen würden das gegenwärtige technologische Niveau stationär fortschreiben. Der CO_2-Düngungseffekt würde nicht berücksichtigt werden, könnte jedoch nachträglich durch eine externe Modellierung einbezogen werden. Durch die doppelte Niveauneutralisierung sind die Modelle besonders für die kombinierte Anwendung mit Wetter- und Klimamodellen geeignet. Systematische Fehler in diesen Modellen würden sich nicht auf die Ertragssimulationen auswirken, da diese sich auf Niveauänderungen und nicht auf das Niveau selbst beziehen.

Über die Aggregation wurden höhere Erklärungswerte erreicht. Durch die Aggregation wurden Ausreißer gemildert, indem die Residuen verringert wurden. Diesen Aggregationseffekt finden auch Woodard und Garcia (2008) und Lobell und Burke (2010). Allerdings diskutieren Garcia et al. (1987), dass bei aggregierten Modellierungsebenen regionale Ertragseinflüsse nicht erfasst und damit der Witterungseffekt unterschätzt wird. Da in unseren Modellen aber nur die Aussageebene und nicht die Modellierungsebene selbst aggregiert wurde, sollte dies nicht relevant sein. Garcia et al. (1987) fordern weiterhin die Verwendung von Betriebserträgen. Diese enthalten jedoch viele betriebsindividuelle Einflüsse (Variation in der Bestandsführung) und erhöhen damit die Unsicherheit bei der Parameterschätzung. Auf Landkreisebene heben sie sich in Deutschland wahrscheinlich größtenteils auf. Die Verwendung von Landkreiserträgen sollte daher ausreichend sein für die Abbildung räumlicher Volatilität. Modellabhängig ließ sich ein Aggregationseffekt zeigen, der beim Übergang von den Landkreisen zu(r) (Sub)Nation(en) zu einer Verbesserung der Modellqualität führte. Je höher die Aggregation war desto größer war der Aggregationseffekt, d.h. die Differenz zwischen dem Mittel der Landkreise und der aggregierten ((sub)nationalen) Modellgütemaße.

Die landkreisindividuellen Parameter der STSMs reproduzierten außergewöhnliche, landkreisindividuelle Ertragsschwankungen vergleichsweise gut. Dies war bei Nutzung von PDMs kaum und bei Verwendung von RCMs nur eingeschränkt möglich. Der Vorteil der STSMs gegenüber den PDMs und RCMs erklärt sich aus der Schätzmethode. Während sich die Parameter der PDMs zwischen den

Landkreisen nicht unterscheiden und die der RCMs tendenziell weniger variieren, gibt es bei den STSM-Parameter eine stärkere Variation (vgl. Beck und Katz, 2007). PDMs und RCMs besitzen Vorteile, wenn die Datenreihen größere Lücken aufweisen. In diesen Fällen kann es bei den STSMs zu größeren Fehlschätzungen kommen. In Abhängigkeit vom Umfang der Datenlücken, kann dann alternativ auf RCMs und PDMs zurückgegriffen werden.

3.6.2 Parametercluster und funktionale Form

Für die STSM-Parameter der Witterungsvariablen zeigten sich deutliche räumliche Cluster. Diese waren nicht deckungsgleich mit den Bundesländern oder Flusseinzugsgebieten. Vielmehr orientierten sie sich an pedosphärischen und topographischen Lagen und spiegelten die Anbaueignung von Winterweizen und Silomais nach Boden- und Klimalagen wieder. Silomais hat als C_4-Pflanze eine höhere Strahlungsnutzungseffizienz als Winterweizen (Monteith, 1977). Letzterer reagiert daher sensitiv auf relative *SRT*-Änderungen. Die räumlichen Cluster der *SRT*-Parameter ließen sich über ähnliche Cluster absoluten Niveaus der *SRT* erklären. Chmielewski und Köhn (2000) stellen zu Winterroggen heraus, dass die Sonnenscheindauer gerade nach der Blütephase einen deutlich positiven Effekt auf den Ertrag hat. Dies deckt sich mit unsern Ergebnissen für Winterweizen. Die Erträge von Winterweizen und Silomais reagierten unterschiedlich sensitiv auf eine Änderung der Wasserversorgung. Silomais reagierte im Osten positiv auf Niederschlagszunahmen in den Monaten Mai bis Juli, da hierdurch seine Jugendentwicklung gefördert wird. Im Alpenvorland zeigten die Parameter nur geringe Ertragseffekte. Dies kann über die ausreichende Wasserversorgung erklärt werden. Für den Winterweizen waren relative *ETP*-Änderungen bedeutender als für Silomais. Von Mai bis Juli ist die Jugendentwicklung des Winterweizens abgeschlossen und dieser befindet sich in der Phase des generativen Wachstums (DWD, 2014). Bei einem *ETP*-Anstieg kommt es auf den wasserlimitierten Standorten (sandige Böden) schnell zur Limitierung von ertragsrelevanten Prozessen. Für Silomais war der *ETP*-Ertragseffekt auf stauwassergefährdeten und flachgründigen Böden (Elbsandsteingebirge, Donauquellgebiet) leicht positiv. Roberts et al. (2012) stellen für die USA in den Monaten vor der Ernte (Juli-August) einen negativen Ertragseffekt durch das *VPD*, für die gesamte Wachstumsperiode aber einen positiven Ertragseffekt durch das *VPD* heraus. Chmielewski und Köhn (2000) zeigen hingegen einen negativen Ertragseffekt für *ETP* während der Winterruhe und einen positiven Effekt von der Blüte bis zur Ernte. Beide berücksichtigen aber zusätzlich noch die Wachstumsgradtage (Berechnung siehe Appendix A.3) bzw. die Temperatur als Variable, die in beiden Studien einen negativen Ertragseffekt hat. Da wir die Wachstumsgradtage nicht berücksichtigten (wegen der hohen Korrelation (0.91^{***}) mit der *ETP*) und unsere *ETP* daher beide Einflüsse abdeckt, waren die Parameter deutlicher negativ. Entsprechend schlussfolgern wir, dass in unseren Parametern der negative Effekt der Wachstumsgradtage enthalten ist. Während der Jugendentwicklung des Winterweizens von November bis April gab es jedoch, analog zum Silomais, einen unmittelbar positiven Effekt von interannuellen Niederschlagsanstiegen auf ertragsbildende Prozesse. Der hohe Niederschlagseffekt im Osten Deutschlands ist auf das geringe absolute Niederschlagsniveau und auf das geringe Wasserspeichervermögen der sandigen Böden in der Region zurückzuführen. Im Alpenvorland wirkte ein Anstieg des Niederschlags negativ. Der negative Niederschlagseinfluss kann auf das hohe Niederschlagsniveau in der Region zurückgeführt werden.

Durch die Cobb-Douglas Produktionsfunktion ist der Ertragseinfluss der Parameter direkt miteinander vergleichbar. Generell zeigten die Niederschlagsparameter bei beiden Kulturen in allen Modellen nur einen sehr geringen Ertragseinfluss. Eine Begründung dafür könnte sein, dass die zu Monaten und Landkreisen aggregierten Niederschlagsvariablen nur begrenzt lokale Niederschlagsextreme oder kurze Trockenperioden erfassen. Da aber auch Trockenjahre wie 2003 von den Modellen abgebildet werden, befinden wir die zeitliche Aggregation der Witterungsvariablen als ausreichend. Bedeutender als die Aggregation sind die Eigenschaften der Variablen und der Methodik des statistischen Schätzverfahrens. Der Ertragseinfluss einer Witterungsvariablen findet nicht nur in der absoluten Größe des Parameterwertes (β_j) seinen Ausdruck, sondern auch in der Ertragsvolatilität, die durch das Produkt von Parameter und exogener Variable errechnet wird (erklärte Varianz von $\beta_j\ x_j$). In unserem Datensatz schwankte der Niederschlag um ±43%. Im Vergleich dazu schwankten *ETP* und *SRT* nur um ±17% und ±8% (Variablen als log erste Differenzen, Zeitraum Mai bis Juli). Durch die hohe Volatilität der Niederschlagsvariable (relative SD) fielen die Niederschlagsparameter daher klein aus. Für die Bewertung des Ertragseinflusses der Witterung ist aber die erklärte Ertragsvolatilität wichtiger als die Parametergröße. Hier zeigte sich, dass die durch den Niederschlag erklärte Ertragsvariation deutlich größer ist als der Ertragseinfluss der Niederschlagsparameter. Dies könnte auch die von Moore und Lobell (2014) gezeigte geringe Ertragswirkung des Niederschlages erklären, da die Autoren sich bei Ihrer Analyse auf die Parameterwerte beschränken.

Zwischen den exemplarisch getesteten Kulturen Winterweizen und Silomais bestanden auf Deutschlandebene nur geringe Unterschiede in der Modellgüte. Die teilweise höheren Erklärungswerte beim Winterweizen auf Subnationsebene ließen sich im Westen von Deutschland über die Volatilität der Erträge erklären. Im Westen von Deutschland war die Varianz der relativen Ertragsänderungen von Winterweizen größer-gleich der von Silomais. Im Osten von Deutschland war hingegen die Varianz der relativen Silomaisertragsänderungen deutlich größer oder zumindest gleich den relativen Winterweizenertragsänderungen. Die höhere Ertragsvolatilität der Silomaiserträge im Osten zeigen, dass Silomais auf trockenen Standorten deutlich sensitiver auf die Witterung reagiert. Dieser Effekt wird noch verstärkt durch die Ausdehnung des Silomaisanbaus in den letzten Jahren auf marginale, sandige Standorte (Krause, 2008). Auf diesen Standorten können winterannuelle Kulturen ungünstige Wachstumsbedingungen über die Ertragskomponenten eher kompensieren (Chmielewski und Köhn, 2000). Umgekehrt ist bei ausreichender Wasserversorgung die Witterungssensitivität von Winterweizen höher und somit Silomais robuster gegenüber der Witterung. Dies konnte gebietsweise im Westen von

Deutschland beobachtet werden. Die Bedeutung der Wasserversorgung von Winterweizen in Deutschland wird ebenfalls von Kersebaum und Nendel (2014) herausgestellt.

Die Einteilung der Witterungsvariablen nach Monaten war eine Annäherung an die zeitlich und räumlich unterschiedlich eintretenden phänologischen Entwicklungsstadien. Dixon et al. (1994) vergleichen die Einteilung der Witterungsvariablen nach phänologischen Entwicklungsperioden gegenüber Kalendermonaten. Sie zeigen, dass eine räumlich und zeitlich differenziere Einteilung nach phänologischen Entwicklungsstadien nur geringe Wirkung auf die Erklärungswerte und Voraussagefähigkeiten von statistischen Modellen hat. Zudem ermöglicht unsere grobe Einteilung die Nutzung der Ertragsmodelle in Kombination mit Klimamodellen und gewährleistet eine weitgehende Unabhängigkeit zwischen den Variablen (Multikollinearität).

3.6.3 Multikollinearität und Verzerrung durch unberücksichtigte Variablen

Da die von uns verwendeten Modelle keine Multikollinearität enthalten, kann ein potenzieller *omitted-variable bias* vernachlässigt werden. Fehlende ertragsrelevante Variablen bedingen dann lediglich eine geringere Erklärungskraft der Modelle. Lobell et al. (2013) zeigen, dass das *VPD* und nicht die Temperatur pflanzenphysiologisch auf den Maisertrag in den USA wirkt. Die in dieser Studie genutzte *ETP* hat gegenüber der von Roberts et al. (2012) und Lobell et al. (2013) genutzten *VPD* den Vorteil, dass sie die Pflanzenbedeckung über den Haude-Faktor berücksichtigt. Reidsma et al. (2007) zeigten, dass Modelle ohne ökonomische Variablen den Witterungseffekt überschätzen. Durch die ökonomischen Proxyvariablen war davon auszugehen, dass die Parameter der Witterungsvariablen unverzerrt sind. Durch die Korrelation der Proxyvariablen untereinander war hier eine Verzerrung der Parameter durch *omitted-variable bias* nicht ausgeschlossen. In der vorliegenden Studie wurden die Parameter der Proxyvariablen nicht weiter verfolgt, da sie den Ertragseinfluss des Proxys enthalten und daher nicht interpretiert werden können (Wooldridge, 2013: 298-300). Unsere Variablenauswahl war pflanzenphysiologisch und produktionstechnisch begründet. Die sehr hohen Erklärungswerte der Modelle sprechen dafür, dass wir die ertragsrelevanten Einflüsse in unseren Modellen berücksichtigen. Modelle, die eine schrittweise Variablenauswahl nutzen, erfassen eventuell aufgrund mangelnder Signifikanz ertragsrelevante Einflüsse nicht und sind daher verzerrt (siehe Appendix A.4).

Die Variablenauswahl erfolgte einheitlich für alle Landkreise. Teilweise wurden die Parameter nicht signifikant verschieden von Null geschätzt (siehe dazu Appendix A.4). Somit fand über die Parameter eine Korrektur der generellen Variablenauswahl statt. Die NSEs der geschätzten und validierten Erträge zeigten aber, dass, trotz teilweise nicht signifikanter Modelle, nicht durchgehend erfüllter OLS-Bedingungen oder Fehlspezifikation, die Modelle robuste Ergebnisse liefern und für Voraussagen geeignet sind.

3.7 Schlussfolgerung

Mit den Modellen können Ertragsabschätzungen für die kurz- und mittelfristige Zukunft bei unterschiedlicher Datenlage durchgeführt werden. Mit zunehmender Vollständigkeit der Datenreihen nimmt die Eignung der Modelle in der Reihenfolge PDM, RCM und STSM zu. Die Modelle können als Entscheidungshilfe bei Investitionen (z.B. in Beregnungstechnik) oder bei der Bepreisung des Risikos in Wetterderivate genutzt werden.

3.8 Literatur

Asseng, S. et al. 2013: Uncertainty in simulating wheat yields under climate change. Nature Climate Change, 3, 827-832.

Bakker, M.M., G. Govers, F. Ewert, M. Rounsevell, R. Jones, 2005: Variability in regional wheat yields as a function of climate, soil and economic variables: Assessing the risk of confounding. Agriculture, Ecosystems & Environment, 110 (3-4), 195-209.

Beck, N., Katz, J.N., 2007: Random Coefficient Models for Time-Series–Cross-Section Data: Monte Carlo Experiments. Political Analysis, 15, 182-195.

Bormann, H.B.D., O. Richter, 1996: Effects of Data Availability on Estimation of Evapotranspiration. Physics and Chemistry of the Earth, 21 (3), 171-175.

Brisson, N., P. Gate, D. Gouache, G. Charmet, F.-X. Oury, F. Huard, 2010: Why are wheat yields stagnating in Europe? A comprehensive data analysis for France. Field Crops Research, 119, 201-212.

Bristow, K.L., G.S. Campbell, 1984: On the relationship between incoming solar radiation and daily maximum and minimum temperature. Agricultural and Forest Meteorology, 31, 159-166.

Butler, E.E., P. Huybers, 2013: Adaptation of US maize to temperature variations. Nature Climate Change, 3, 68-72.

Castellvi, F., P.J. Perez, J.M. Villar, J.I. Rosell, 1996: Analysis of methods for estimating vapor pressure deficits and relative humidity. Agricultural and Forest Meteorology, 82, 29-45.

Challinor, A.J., J. Watson, D.B. Lobell, S.M. Howden, D.R. Smith, N. Chhetri, 2014: A meta-analysis of crop yield under climate change and adaptation. Nature Climate Change, 4, 287-291.

Chmielewski, F.-M., A. Müller, E. Bruns, 2004: Climate changes and trends in phenology of fruit trees and field crops in Germany, 1961-2000. Agricultural and Forest Meteorology, 121, 69-78.

Chmielewski, F., W. Köhn, 2000: Impact of weather on yield components of winter rye over 30 years. Agricultural and Forest Meteorology, 102, 253-261.

Croissant, Y., G. Millo, 2008: Panel Data Econometrics in R: The plm Package. Journal of Statistical Software, 27 (2), 1-43.

Dielman, T.E., 1983: Pooled Cross-Sectional and Time Series Data: A Survey of Current Statistical Methodology. The American Statistician, 37 (2), 111-122.

Dixon, B.L., S.E. Hollinger, P. Garcia, V. Tirupattur, 1994: Estimating Corn Yield Response Models to Predict Impacts of Climate Change. Journal of Agricultural and Resource Economics, 19 (1), 58-68.

Doll, J.P., 1967: An Analytical Technique for Estimating Weather Indexes from Meteorological Measurements. Journal of Farm Economics, 49 (1), 79-88.

Donatelli, M., G. Bellocchi, L. Carlini, 2006: Sharing knowledge via software components: Models on reference evapotranspiration. European Journal of Agronomy, 24 (2), 186-192.

DVWK, 1996: Ermittlung der Verdunstung von Land- und Wasserflächen. DVWK-Merkblätter zur Wasserwirtschaft. Wirtschafts- und Verlagsgesellschaft Gas und Wasser mbH, Bonn, 240 S.

DWD, 2014: Aktueller Stand der Phänologie in Deutschland, 1992-2013. Deutscher Wetterdienst, Retrieved from: http://www.dwd.de/phaenologie (12.04.2015).

Garcia, P., S.E. Offutt, M. Pinar, S.A. Changing, 1987: Corn Yield Behavior - Effects of Technological Advance and Weather-Conditions. Journal of climate an applied meteorology, 26, 1092-1102.

Iizumi, T., H. Sakuma, M. Yokozawa, J.-J. Luo, A.J. Challinor, M.E. Brown, G. Sakurai, T. Yamagata, 2013: Prediction of seasonal climate-induced variations in global food production. Nature Climate Change, 3, 904-908.

Kaufmann, R.K., S.E. Snell, 1997: A Biophysical Model of Corn Yield: Integrating Climatic and Social Determinants. American Journal of Agricultural Economics, 79 (1), 178-190.

Kersebaum, K.C., C. Nendel, 2014: Site-specific impacts of climate change on wheat production across regions of Germany using different CO^2 response functions. European Journal of Agronomy, 52, 22-32.

Krause, J., 2008: A Bayesian approach to German agricultural yield expectations. Agricultural Finance Review, 68, 9-23.

Krause, P., D.P. Boyle, F. Bäse, 2005: Comparison of different efficiency criteria for hydrological model assessment. Advances in Geosciences, 5, 89-97.

Lee, B.-H., P. Kenkel, B.W. Brorsen, 2013: Pre-harvest forecasting of county wheat yield and wheat quality using weather information. Agricultural and Forest Meteorology, 168, 26-35.

Lobell, D.B., 2010: Crop Responses to Climate: Time-Series Models. In: D. Lobell und M. Burke (Editors), Climate Change and Food Security. Advances in Global Change Research. Springer Netherlands, S. 85-98.

Lobell, D.B., 2013: Errors in climate datasets and their effects on statistical crop models. Agricultural and Forest Meteorology, 170, 58-66.

Lobell, D.B., G.P. Asner, 2003: Climate and Management Contributions to Recent Trends in U.S. Agricultural Yields. Science, 299 (5609), 1032.

Lobell, D.B., M.B. Burke, 2010: On the use of statistical models to predict crop yield responses to climate change. Agricultural and Forest Meteorology, 150 (11), 1443-1452.

Lobell, D.B., G.L. Hammer, G. McLean, C. Messina, M.J. Roberts, W. Schlenker, 2013: The critical role of extreme heat for maize production in the United States. Nature Climate Change, 3 (5), 497-501.
Lobell, D.B., W. Schlenker, J. Costa-Roberts, 2011: Climate trends and global crop production since 1980. Science, 333 (6042), 616-620.
Monteith, J.L., 1977: Climate and the efficiency of crop production in Britain. Philosophical Transactions of the Royal Society of London, 281, 277-294.
Moore, F.C., D.B. Lobell, 2014: Adaptation potential of European agriculture in response to climate change. Nature Climate Change, 4, 610-614.
Nendel, C., R. Wieland, W. Mirschel, X. Specka, C. Guddat, K.C. Kersebaum, 2013: Simulating regional winter wheat yields using input data of different spatial resolution. Field Crops Research, 145, 67-77.
Oury, B., 1965: Allowing for Weather in Crop Production Model Building. Journal of Farm Economics, 47 (2), 270-283.
Palosuo, T., K.C. Kersebaum, C. Angulo, P. Hlavinka, M. Moriondo, J.E. Olesen, R.H. Patil, F. Ruget, C. Rumbaur, J. Takáč, M. Trnka, M. Bindi, B. Çaldağ, F. Ewert, R. Ferrise, W. Mirschel, L. Şaylan, B. Šiška, R. Rötter, 2011: Simulation of winter wheat yield and its variability in different climates of Europe: A comparison of eight crop growth models. European Journal of Agronomy, 35 (3), 103-114.
Reidsma, P., F. Ewert, A.O. Lansink, 2007: Analysis of farm performance in Europe under different climatic and management conditions to improve understanding of adaptive capacity. Climatic Change, 84 (3-4), 403-422.
Roberts, M.J., W. Schlenker, J. Eyer, 2012: Agronomic weather measures in econometric models of crop yield with implications for climate change. American Journal of Agricultural Economics, 95 (2), 236-243.
Schlenker, W., M.J. Roberts, 2009: Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. Proceedings of the National Academy of Sciences of the United States of America, 106 (37), 15594-15598.
Schrödter, H., 1985: Verdunstung: Anwendungsorientierte Meßverfahren und Bestimmungsmethoden. Springer, 204 S.
Shaw, L.H., 1964: The Effect of Weather on Agricultural Output: A Look at Methodology. Journal of Farm Economics, 46 (1), 218-230.
Sonntag, D., 1990: Important new Values of the Physical Constants of 1986, Vapour Pressure Formulations based on ITS-90, and Psychrometer Formulae. Meteorologische Zeitschrift, 4 (5), 340-344.
Statistisches Bundesamt, 2013: Ackerland nach Hauptfruchtgruppen und Fruchtarten, Land- & Forstwirtschaft, Fischerei - Feldfrüchte und Grünland.
Tannura, M.A., S.H. Irwin, D.L. Good, 2008: Weather, Technology, and Corn and Soybean Yields in the U.S. Corn Belt. Marketing and Outlook Research Report, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, 1-127.
Woodard, J.D., P. Garcia, 2008: Weather Derivatives, Spatial Aggregation, and Systemic Risk: Implications for Reinsurance Hedging. Journal of Agricultural and Resource Economics, 33 (1), 34-51.
Wooldridge, J.M., 2013: Introductory Econometrics. A Modern Approach. South Western Cengage Learning, 868 S.
You, L., M.W. Rosegrant, S. Wood, D. Sun, 2009: Impact of growing season temperature on wheat productivity in China. Agricultural and Forest Meteorology, 149 (6-7), 1009-1014.

4 Statistical regression models for assessing weather impacts on crop yields – A validation study for winter wheat and silage maize in Germany

Christoph Gornott[1*] and Frank Wechsung[1]

[1] Potsdam Institute for Climate Impact Research (PIK)

[*] Corresponding author

4.1 Abstract

For agriculture in Germany and generally all around the world, yield variability due to uncertain weather conditions represents an increasing production risk. Regional assessments of future yield changes can help farmers to cope with this risk. For Germany's two most important crops winter wheat (*Triticum aestivum* L.) and silage maize (*Zea mays* L.), we investigate three regression models estimating relative weather impacts on relative crop yield changes: the separate time series model (STSM), the panel data model (PDM) and the random coefficient model (RCM). These regression models use the Cobb–Douglas function to capture weather and non-weather impacts on yields (e.g. changing prices or inventory management). The yield influencing weather impacts contain the potential growth and stress factors during vegetative and reproductive plant development. The models are estimated and validated at the county scale. To improve the robustness and goodness of fit, the models are aggregated at the scale of German federal states, river basins and at the national scale. The observed yield changes are satisfactorily reproduced by all models for all aggregated scales (measured by the Nash–Sutcliffe efficiency (NSE)). According to their NSE values, the methodically simple STSMs reproduce extreme yield changes better (0.85) than the RCMs (0.79) and PDMs (0.72) at the national scale. This order can be also found across all scales when considering the models' goodness of fit. Generally, spatial aggregation increases the goodness of fit by +0.16 for federal states and river basins and by +0.29 for entire Germany compared to the county scale. The mean NSE increase is lowest for STSMs (+0.11), followed by RCMs (+0.13) and PDMs (+0.25) for federal states and river basins, which is opposite to the goodness of fit order. The model parameters show clear spatial patterns, which reflect regional differences of weather and soil. Within its methodological limits, our approach can directly be combined with the output of climate models and is suitable for assessing short- and medium-term yield effects for the current agronomic practice. It requires neither bias correction of the climate variables nor explicit modeling of crop yield trends.

Keywords:

Statistical crop yield models, Weather impacts, Yield changes, Winter wheat, Silage maize, Germany

4.2 Introduction

4.2.1 Statistical crop models for yield assessments

Crop yield assessments for upcoming climate anomalies or altered weather conditions are of general interest for farmers, traders (e.g. grain mills, retailers), insurance companies, and policy makers. Statistical models (Ray et al., 2015; Iizumi et al., 2013; Mueller et al., 2012; Roberts et al., 2012; Schlenker and Roberts, 2009) and process based models (Asseng et al., 2013; Angulo et al., 2013; Palosuo et al., 2011) are model types for such assessments. Both model types are parametrized for past weather records. For future projections, they need weather records from climate simulation models. These climate models very often require a bias correction of the simulated output before using them for yield projections (Lobell, 2013).

Process based crop models may not include all weather related effects on crop yields. There are many yield effects, which simply cannot be captured in process based models, because of limited spatial information about these effects. Examples are weather-triggered effects on agronomic adaptation (irrigation, crop varieties, agronomic techniques) or on pests, weeds, and diseases (Mueller et al., 2012). These weather-triggered effects can be collinear with the weather variables. Since crop yields also contain weather-triggered effects, statistical crop yield models estimate in their parameter values not only the sole, but also the triggered effect of the weather variable. Process based models do not capture these weather-triggered effects as long as they are not explicitly embedded in the models (Estes et al., 2013; Lobell and Burke, 2010). In the assessment of farm level yield effects, this is an important disadvantage of process based models in comparison to statistical models.

Statistical yield models also allow relating inter-annual yield and yield factor changes (i.e. first order temporal ratios) instead of absolute values to each other (You et al., 2009; Lobell, 2007; Lobell and Asner, 2003). Considering changes instead of absolute values eliminates the trend of the variables and it allows neglecting systematic biases for exogenous variables for example when using simulated climate data from circulation models (Lobell, 2013). However, the neglected absolute level by using changes ignores a possible level dependency of yield and climate conditions. This limits the suitability to climate change assessments for changes within the range of recent climate variability. For yield projections beyond the yield variability of the dataset used for model estimation, process based models might be more appropriate (Rötter et al., 2011). At least, process based models should complement the statistical assessments under such circumstances.

The impact of weather on crop yields can be subdivided into two variable groups: variables that primarily determine potential growth and those that can be related to stress influences. The distinction is not disjunctive, overlaps might exist. We focus on the main influences that can contribute to a statistical explanation of the yield variability. The potential yield is determined mainly by the incoming solar radiation (Monteith, 1977; Long et al., 2006). The best usage of this incoming solar radiation requires an optimal mix of agronomic measures to establish the crop, to supply the necessary nutrients and water, and to keep biotic stress factors under control. Any divergence from this optimal mix will result in stress that reduces the potential yield. For these potential stress factors, we distinguish two groups: weather and management driven stress factors.

Among all possible weather driven stress factors, we hypothesize water stress as the most relevant stress factor for German winter wheat and silage maize yields (Wessolek and Asseng, 2006; Kersebaum and Nendel, 2014; Wolf and Diepen, 1994). Other possible influences, like temperature stress, might also exist in single years (Rötter and van de Geijn, 1999; Lobell et al., 2013), but are less generally associable with German weather conditions. Management driven stress factors, like the crop variety, fertilizer, plant protection, and machinery, are reflected in the mean yield level and the yield trend. However, there are also economic conditions, e.g., statutory set-aside quotas or renewable energy subsidies for biogas and biodiesel, which influence the annual yield variability (Krause, 2008; Bakker et al., 2005). We use the fertilizer price and the acreage of the respective crops as proxy variables to control the economic yield impacts in the models. The fertilizer price represents the varying profitability of production factor inputs (e.g. seeds, plant protection, fuel, and fertilizer) and may directly affect the yield variability. The acreage of winter wheat and silage maize represents changes in the Common Agricultural Policy (CAP) of the European Union. An expanded acreage might generally suppress the yield level of both crops due to the inclusion of marginal productive land.

4.2.2 Modeling approach

In our approach, we follow the modeling concept introduced by Wechsung et al. (2008) and the validation scheme of Gornott and Wechsung (2015), who expanded the concept by two other statistical approaches. A level neutralizing transformation is applied for all variables, i.e., the crop yield, the weather-related and the non- weather-related variables. We utilize first order ratios $y'_t = \frac{y_t}{y_{t-1}}$ and $x'_t = \frac{x_t}{x_{t-1}}$, for the years $t = 2, \dots, M$ of the endogenous variable crop yield y_t and the exogenous weather and non- weather variables x_t. As functional form, we use the Cobb–Douglas function analogous to Oury (1965). The function is proven in both economic (You et al., 2009) and agronomic applications (Lee et al., 2013) and considers yield impacts arising from substitution and interaction between the exogenous variables. Due to the linearization of the Cobb–Douglas function, the first order ratios are transformed to logarithmic first order ratios of yields and yield-factors, hereafter expressed as yield and factor changes. These changes allow an intercomparison of the effects of different variables.

We test three alternative ways to incorporate the spatial heterogeneity of the relationships between yield changes and yield factor changes: by separately estimated time series models (STSMs), panel data models (PDMs), and random coefficient models (RCMs). All three approaches refer to a spatial dataset consisting of N discrete subunits and M years. In our case, the subunits are German counties within a federal state, river basin, or Germany as a whole. The methodically simple STSMs are estimated independently for the N subunits resulting in N parameter sets (Butler and Huybers, 2013; Lobell and Burke, 2010). In contrast, PDMs capture directly the temporal and spatial variability by one parameter set for all of the considered N subunits (You et al., 2009). RCMs can be ranked between PDMs and STSMs. They allow individual parameter variations per subunit and a parameter set for the entire unit (Reidsma et al., 2007). The results of the estimations will be presented and evaluated at two scales: the original spatial data scale, i.e., the German county yields, and the aggregated data scale, i.e. federal states, river basins, and entire Germany. Due to the aggregation, county- and farm-individual influences are largely averaged out, which might have biased the model results otherwise (Woodard and Garcia, 2008).

We restricted the temporal and spatial resolution of all variables to a division, which is accessible for climate simulations. The model results are evaluated at a larger scale than the estimation scale. Thus, we make explicit use of spatial aggregation effects. We test and apply the approach in respect to its possible suitability for fast impact assessment of seasonal- and medium-term projections (up to 30 years) from climate models. The approach is conducted for winter wheat and silage maize, because these are the major winter and summer annual crops in Germany.

4.3 Materials and methods

4.3.1 Data

We use a spatial dataset of German crop yields per county for winter wheat and silage maize from 1991 to 2010. The dataset is supplied by the Statistical Offices of the Federation and the Länder (2013b). Weather data are available for the same period from 1,218 German weather stations (DWD, 2011). The data are averaged per county to match the spatial resolution of the crop yield data. The total acreage of winter wheat and silage maize is taken from the datasets of the Statistical Offices of the Federation and the Länder (2013a) [1991–2008] and the Federal Statistical Office (2013) [2008–2010]. The fertilizer price index is published by the Statistical Offices of the Federation and the Länder (2013c). Ideally, all variables would be estimated at the county scale. However, the economic variables are only available on a national scale, so we applied the national values to all counties. A detailed description of the data is contained in the supplemental information (SI) S.1.

4.3.2 Model approach

4.3.2.1 Basic function

The Cobb–Douglas function is used as the basic function in all statistical models (Eq. 1). The function relates inter-annual changes of crop yield (y'_t) to J weather (x'_{jt}) and K economic variables (x'_{kt}). Statistical models often have the disadvantage that the parameter values are not easily accessible for an interpretation. In our approach, the parameter values of the Cobb–Douglas function are directly comparable per and across crops and spatial sites as relative yield effects by a relative increase of the exogenous variables (Wooldridge, 2013, p. 351-354).

$$y'_t = \beta_0 \prod_{j=1}^{J} \left(x'_{j\,t}\right)^{\beta_j} \prod_{k=1}^{K} (x'_{k\,t})^{\beta_k}\text{, with } j = 1, \dots, J \text{ and } k = 1, \dots, K \tag{1}$$

β_0 - mean annual changes of y'_t,

β_j, β_k - partial relative change of y' per unit change of x'_{jt} and x'_{kt}, respectively.

4.3.2.2 Regression models

The basic function (Eq. 1) can be linearized by logarithm. The variables are transformed to linear terms and the function is expanded by an error term u_t to become accessible for regression analysis. The spatial yield variability within Germany, German federal states and river basins is addressed using three alternative regression models: STSMs, PDMs, and RCMs.

STSMs (Eq. 2) separately consider the individual yield changes at the *N* subunits, by estimating independently a series of *N* models (Dielman, 1983).

$$\log\, y'_{it} = \log \beta_{i0} + \sum_{j=1}^{J} \beta_{ij} \log\, x'_{ijt} + \sum_{k=1}^{K} \beta_{ik} \log\, x'_{kt} + \log\, u'_{it}, \qquad i = 1, \dots, N \tag{2}$$

For our approach, the values of the x'_{kt} variables do not vary by the index *i* as all other variables, because county individual data of the economic variables are not available. Therefore only national values of those variables are related to the county individual crop yields.

Unlike STSMs, PDMs (Eq. 3) estimate directly one parameter set for all *N* subunits. The parameter values (β_j, β_k) do not vary among the *N* subunits as for STSMs. PDMs may still capture county individual, time-invariant effects (e.g. soil productivity and farm size effects) due to the normalizing effect of the county-wise first difference transformation. These effects, which are contained in the mean yield level, are eliminated by the first order transformation before model estimation (Wooldridge, 2013; Croissant and Millo, 2008). When this transformation is reversed after calculating absolute crop yields, then the spatial differences in the mean yield level re-appear.

$$\log\, y'_{it} = \log \beta_0 + \sum_{j=1}^{J} \beta_j \log\, x'_{ijt} + \sum_{k=1}^{K} \beta_k \log\, x'_{kt} + \log\, u'_{it} \tag{3}$$

RCMs (Eq. 4) may be ranked between STSMs and PDMs. They contain both one parameter set, which is valid for all subunits ($\beta_0, \beta_j, \beta_k$) and county individual parameter variations (b_{i0}, b_{ij}, b_{ik}). The county-individual impact β_i results from $\beta_i = \beta + b_i$. Since the parameters β and b_i are dependent on each other, the model cannot be estimated by the ordinary least squares method (OLS). Instead, our RCMs are estimated by the restricted maximum likelihood method (REML) analogous to Reidsma et al. (2007).

$$\log\, y'_{it} = \log(\beta_0 + b_{i0}) + \sum_{j=1}^{J} \left(\beta_j + b_{ij}\right) \log\, x'_{ijt} + \sum_{k=1}^{K} (\beta_k + b_{ik}) \log\, x'_{kt} + \log\, u'_{it} \tag{4}$$

4.3.2.3 Aggregation of the model results

The estimated and measured *N* individual yield changes per county are averaged to the arithmetic mean (Eq. 5) for the aggregation scale (i.e. nation, federal states, and river basins) in hindsight. We did not aggregate the exogenous variables, because this would lead to information losses due to a reduced variability during the estimation.

$$\overline{\log\, y'_t} = N^{-1} \sum_{i=1}^{N} \log\, y'_{it} \tag{5}$$

4.3.3 Exogenous variables

The selection of variables aims to capture major weather and economic influences on the crop yield variability. The variables are selected according to their plant physiological impact. Across most of the subunits, but not necessarily in all, the variables are expected to be significant (see SI S.2).

4.3.3.1 Weather variables

The temporal variable division is based on an aggregated view of the plant growth process. Chmielewski and Köhn (2000) distinguish between five phenological development periods in their yield component analysis for winter rye. Butler and Huybers (2015) use for their statistical yield models four phenological development periods, while Moore and Lobell (2014) and You et al. (2009) did not divide the growing period. Dixon et al. (1994) divide the growing period by calendar months and phenological phases. They show that the division by calendar months leads to similar results in comparison to phenological phases. We distinguish two phases: the vegetative and reproductive development. The daily values of the weather variables are separately summed by calendar month over the vegetative and reproductive sections of the winter wheat and silage maize growing period. The vegetative development for winter wheat approximately lasts from November (of the planting year) to April (of the harvest year) [hereafter Nov–Apr]. For silage maize, the vegetative development has an approximate duration from May to July [May–Jul]. The reproductive development of winter wheat fulfills between May and July and that of silage maize between August and October [Aug–Oct] (DWD, 2015).

The selection of weather variables for the model is complicated by the problem of multicollinearity, i.e. a correlation between the exogenous variables. Multicollinearity leads to less precise estimates and large standard errors of the parameters. Independence among different growth variables and coverage of potential growth and stress factors is thought to be achieved by the variables: temperature normalized solar radiation (*SRT*) for the growth potential; precipitation (*PREC*) and potential evapotranspiration (*ETP*) for the water supply and the atmospheric water demand. The *SRT* is used instead of the solar radiation (R_S) to minimize the possibility of multicollinearity with the other variables. The daily *SRT* [J °C^{-1} cm^{-2}] is calculated by Eq. 6. To avoid division by zero, the temperature value is increased by 20 similar to the correction of the de Martonne aridity index (Oury, 1965).

$$SRT = \frac{R_S}{T_{\text{avg}} + 20}, \qquad \text{with} \tag{6}$$

R_S - daily solar radiation sum [J cm^{-2}],
T_{avg} - daily average temperature [°C].

The daily *ETP* [mm] is calculated following Haude (Eq. 7). *ETP* depends on the vapor pressure deficit and an empirical correction factor, the Haude factor f_H (Schrödter, 1985; Haude, 1955). For f_H, we use the arithmetic average of the values for wheat, maize, and grassland for each calendar month (see SI Tab. S.3). This considers respective characteristics of the modeled crops and is available for the calculations of all relevant months (we added the grassland values, because the values for wheat and maize are not available for the entire growing season). The vapor pressure deficit is calculated using the Magnus formula (Sonntag, 1990) by the maximum temperature (T_{max}) and the minimum temperature (T_{min}) instead of dew point temperature. Thereby, we follow Castellvi et al. (1997) and Castellvi et al. (1996) who suggested the replacement of dew point temperature by T_{min} in the calculation of the ETP.

$$ETP = f_H\ 6.11 \left(\exp\left(\frac{17.269\ T_{\max}}{237.3 + T_{\max}}\right) - \exp\left(\frac{17.269\ T_{\min}}{237.3 + T_{\min}}\right)\right), \qquad \text{with} \tag{7}$$

f_H - Haude factor,
$T_{\max}$/ $T_{\min}$ - daily maximum/ minimum temperature.

4.3.3.2 Economic variables

Kaufmann and Snell (1997) argue that an omitted-variable bias may occur in the case of unconsidered yield-related (economic) variables in statistical models. Accordingly, we considered fertilizer price and acreage of the respective crops as economic proxy variables to control the economic yield impacts in the models. The mean annual fertilizer price serves as proxy for a set of input prices in the winter wheat and silage maize production. Price fluctuations within one year are averaged out. The effects of pre-contracts on fertilizer, which might lead to time lag effects, are neglected in our variable setting. The available data does not allow a quantification of those lag-effects yet.

Finally, our winter wheat models contain the variables: *ETP* (Nov–Apr), *ETP* (May–Jul), *SRT* (May–Jul), *PREC* (Nov–Apr), *PREC* (May–Jul), *fertilizer price*, and *acreage of winter wheat*. The silage maize models contain the variables: *ETP* (May–Jul), *ETP* (Aug–Oct), *SRT* (May–Jul), *PREC* (May–Jul), *PREC* (Aug–Oct), *fertilizer price*, and *acreage of silage maize*.

4.3.4 Model fit and robustness

The models should be able to reproduce both the mean level and the variability of the measured yield changes. Both characteristics are assessed by calculating the root-mean-square error (RMSE), the coefficient of determination (R^2), and the Nash–Sutcliffe efficiency coefficient (NSE) (see SI S.3 for the calculation). The RMSE captures the deviation from the mean level. The R^2 measure the reproduction of the variability. The NSE is a combined indicator for the mean model bias and the variability. It is particularly suitable for out-of-sample cross validations (Chipanshi et al., 2015; Krause et al., 2005).

The robustness of the three suggested yield models (STSM, PDM, and RCM) is assessed by running an out-of-sample cross-validation. For each year *t*, the dataset is subsequently reduced by all values of a year *t*. For this reduced dataset, we estimate the model parameters and use these parameters to calculate the yield changes of the removed years. To more rigorously check the robustness of our approach, we expanded the validation process by removing the values of further four randomly selected years, in addition to the year *t*. We refer the one-year validation simply as *validation*, while the five-year validation is called *expanded validation* hereafter.

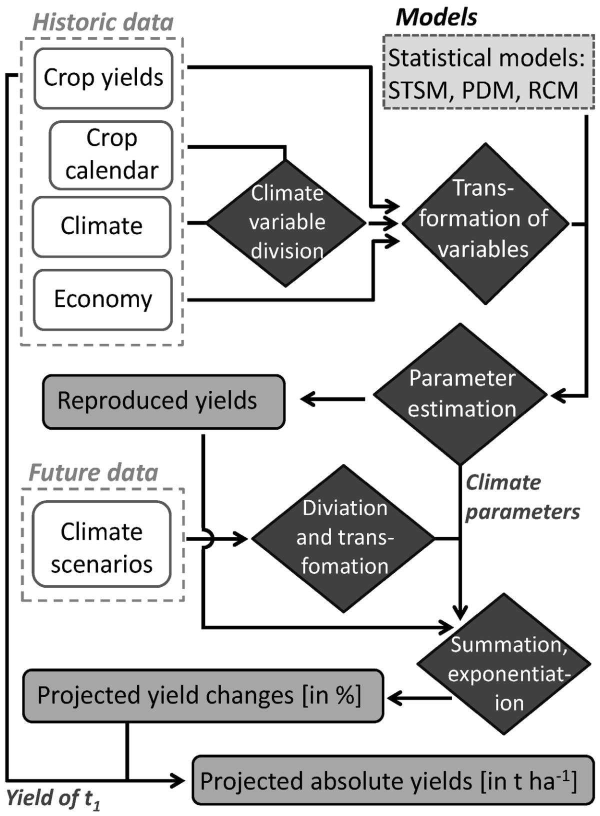
The permissibility of the OLS estimator for STSMs and PDMs is statistically tested using several tests described by Croissant and Millo (2008) and Wooldridge (2013). The regression equation specification error test (RESET) allows an evaluation whether quadratic variables are missed in the models. The Lagrange multiplier test according to Breusch–Pagan (LM) is used to examine the spatial independency (heterogeneity) of the data and justifies a spatial regression approach. Otherwise, averaging across all counties would be sufficient. The Breusch–Godfrey test is applied to assess autocorrelation and the Breusch–Pagan test is used to test for heteroskedasticity. The normal distribution of residuals is tested using the Shapiro–Wilk test. For the RCMs, there exist several criteria for the evaluation of the model goodness of fit (Reidsma et al., 2007). We use the relative criterion NSE, because it is suitable for both OLS and REML (non OLS) conditions (Reidsma et al., 2007; Chipanshi et al., 2015). A description of the applied software is given in the SI S.4.

4.3.5 Model application for yield projection

The calculation of the relative yield changes between the reference period (utilized for the estimation of the years *t*) and a projection period of the years t^* (with $t^* = M + 1, \dots, P$) does not need an explicit specification of the basic yield level y_{t_1}. Nevertheless, it should be kept in mind that any change is related to the agronomic level of that reference. Absolute crop yields for a last year P of the projection period are calculated combining the basic yield level y_1, the yield changes during the reference period, and the projected yield changes of P following Eq. 8:

$$y_P = y_{t_1} \exp\left(\sum_{t=1}^{M} \log y'_t + \sum_{t^*=M+1}^{P} \log y'_{t^*}\right) \tag{8}$$

The concrete application is out of the scope of this paper. Here, we focus on the validation of the modeling approach. In Fig. 1a and 1b we present a schematic and algorithmic description that can be followed during applications nevertheless.



1) Estimation of the weather (β_j) and economic parameters (β_k) by Eq. 2, 3, or 4.
2) Transformation of future climate records to climate changes: $\log\ x'_{jt^*}$.
3) Multiplication and summation of the estimated β_j and the projected climate changes: $y'_{t^*} = \sum_{j=1}^{J} \beta_j \log\ x'_{jt^*}$.

Note: The intercept β_0 and the economic variables are not utilized for the projection.

4) Summation and exponentiation of the past and projected yield changes: $y'_P = \exp(\sum_{t=1}^{M} y'_t + \sum_{t^*=M+1}^{P} y'_{t^*})$.

Note: The calculation until step 4 results in a change factor that can be also converted into percentage yield changes (e.g. 1.10 → +10%).

5) Multiplication of the yield changes with the first yield: y_{t_1} (e.g. 7t ha^{-1}).

Fig. 1: (a) Workflow when using our modeling framework for projecting crop yield impacts due to climate changes and (b) steps of model building and application for accessing weather and climate impacts.

4.4 Results

4.4.1 Goodness of fit

STSMs, PDMs, and RCMs are able to reproduce the measured winter wheat and silage maize yield changes at the aggregated scale for the estimations and the validations (examples shown for STSMs, RCMs, and PDMs in Fig. 2). For both crops and all models, a decrease in NSE by approximately 0.25 is common when comparing estimations with validations. The NSE decreases approximately 0.38 in the expanded validation (not shown).

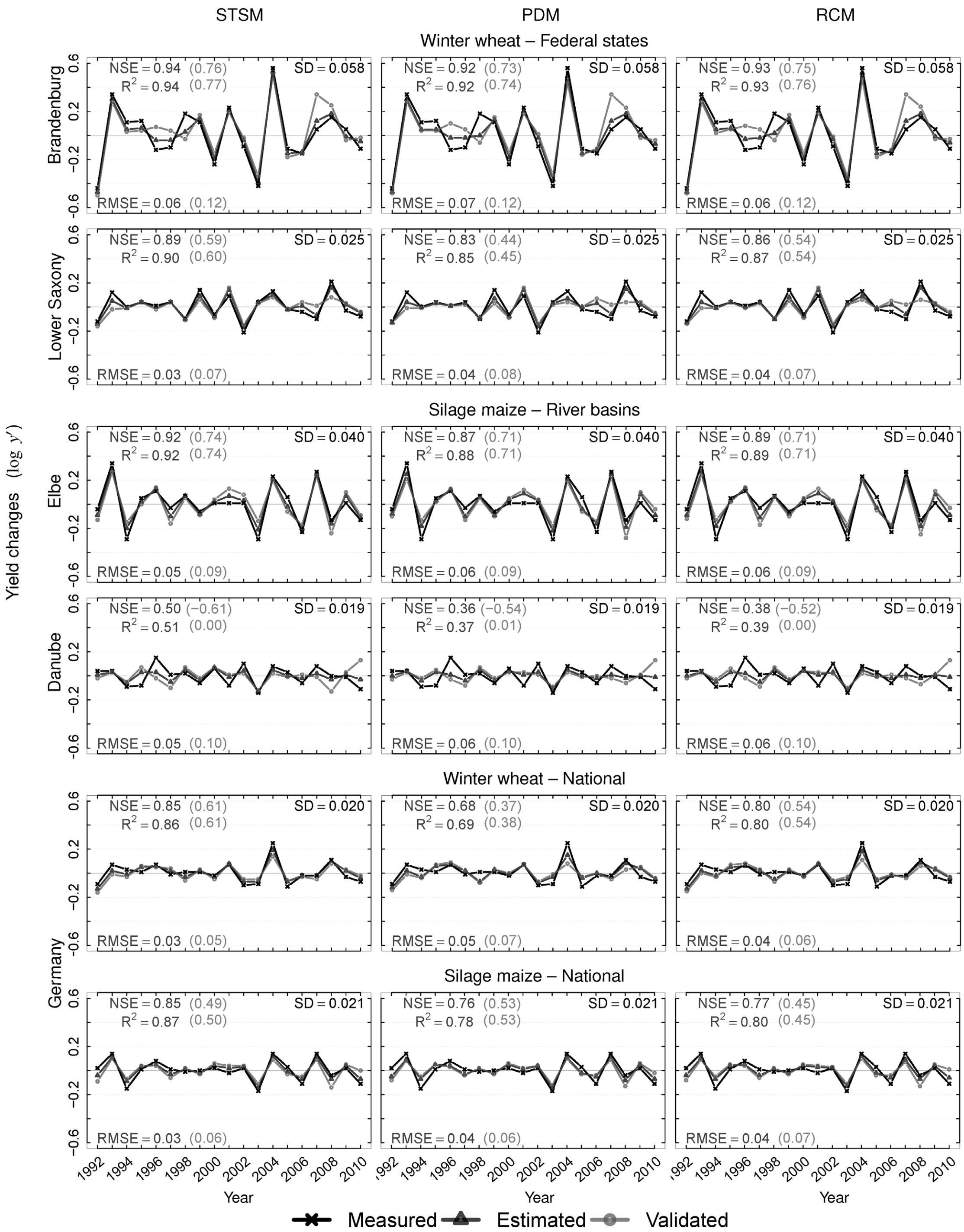


Fig. 2: Time series of measured, estimated, and validated crop yield changes of winter wheat for two German federal states, of silage maize for two river basin and for winter wheat and silage maize, respectively, at the national level using STSMs (left) , PDMs (center), and RCMs (right). The black values for NSE, R^2, and RMSE relate to the model estimation. The gray values in parenthesis characterize the model performance during validation. The SD values are the standard deviation of the measured yields.

Tab 1: The NSE measure for the winter wheat and silage maize crop yield models, the different model types (STSM, PDM, RCM), and aggregation scales (federal states, river basins, national). The arithmetic averages of federal states (FS_{Avg}) and river basins (RB_{Avg}) are given below the aggregation scales. The acronyms are: SH: Schleswig-Holstein, LS: Lower Saxony, NRW: North Rhine-Westphalia, HE: Hesse, RP: Rhineland-Palatinate, BW: Baden-Württemberg, BA: Bavaria, SL: Saarland, BB: Brandenburg, MWP: Mecklenburg-Western Pomerania, SN: Saxony, SA: Saxony-Anhalt, TH: Thuringia, DAN: Danube, ST: Schlei/ Trave, WP: Warnow/ Peene, and GER: Germany.

	Winter wheat			Silage maize		
Unit	STSM	PDM	RCM	STSM	PDM	RCM
Federal states (FS)						
SH	0.75	0.67	0.70	0.60	0.52	0.52
LS	0.89	0.83	0.86	0.78	0.65	0.71
NRW	0.84	0.80	0.81	0.65	0.53	0.57
HE	0.62	0.58	0.58	0.67	0.59	0.64
RP	0.70	0.64	0.66	0.61	0.59	0.58
BW	0.76	0.65	0.69	0.66	0.54	0.59
BA	0.72	0.55	0.64	0.67	0.51	0.58
SL	0.69	0.67	0.68	0.77	0.76	0.77
BB	0.94	0.92	0.93	0.94	0.93	0.93
MWP	0.95	0.91	0.94	0.93	0.90	0.92
SN	0.85	0.73	0.79	0.72	0.58	0.67
SA	0.89	0.86	0.89	0.93	0.92	0.92
TH	0.67	0.63	0.65	0.81	0.78	0.79
FSAvg	0.79	0.73	0.76	0.75	0.68	0.71
River basins (RB)						
Eider	0.75	0.72	0.75	0.43	0.33	0.37
ST	0.69	0.64	0.67	0.54	0.50	0.52
Elbe	0.91	0.81	0.88	0.92	0.87	0.89
Weser	0.88	0.82	0.85	0.84	0.78	0.77
Ems	0.85	0.78	0.83	0.57	0.34	0.50
Rhine	0.78	0.68	0.73	0.76	0.67	0.69
Maas	0.69	0.68	0.68	0.75	0.74	0.75
DAN	0.72	0.62	0.66	0.50	0.36	0.38
WP	0.94	0.91	0.94	0.93	0.90	0.92
Oder	0.75	0.77	0.77	0.93	0.91	0.92
RSAvg	0.80	0.74	0.78	0.72	0.64	0.67
National						
GER	0.85	0.68	0.80	0.85	0.76	0.77

For all models, crops, and aggregation scales, the goodness of fit (measured by the NSE for the aggregated scales) is shown in Tab. 1. Generally, the NSE decreases from STSM over RCM to PDM for both winter wheat and silage maize in the estimations. For both crops, there exists a significant Pearson correlation ($^{***}p \leq 0.01$, $^{**}p \leq 0.05$, $^{*}p \leq 0.1$, $p > 0.1$) between yield variability (i.e., standard deviation (SD)) and goodness of fit (i.e., NSE). However, this correlation is stronger for silage maize (0.86^{***}) than for winter wheat (0.66^{**}) on the federal state scale. This difference is also illustrated by the fact that the SDs of both crops are strongly correlated (0.93^{***}), but the corresponding NSEs correlate only by 0.64^{**}.

Differences between models exist in their reproduction of extreme inter-annual yield changes in single counties (Fig. 3). In the federal state of Brandenburg, for example, the STSMs (NSE: 0.84) reproduce extreme county yield changes (see county LOS: Oder-Spree) considerably better (light yellow) than RCMs (NSE: 0.81) and PDMs (NSE: 0.70) (indicated by the blue, red, and orange color).

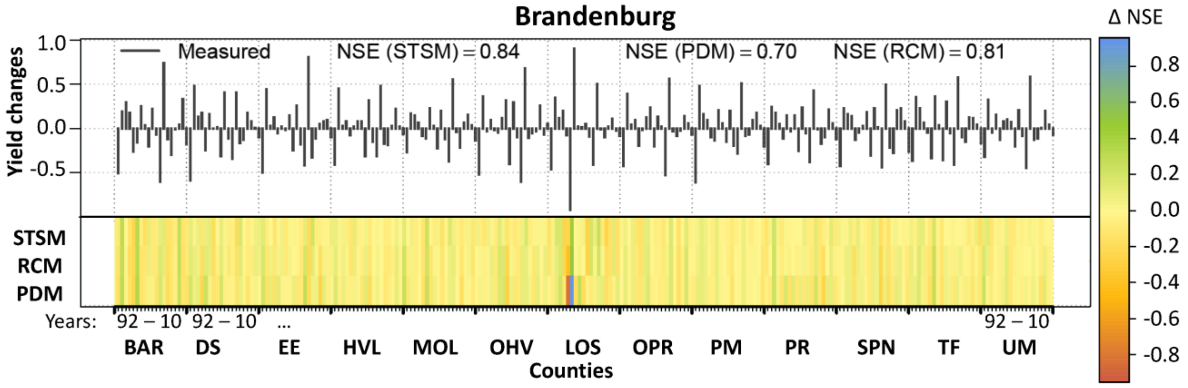


Fig. 3: Time series of measured winter wheat yield changes (bars) and the difference ($\Delta \log y' = \log y'_{measured} - \log y'_{estimated}$) between measured and estimated winter wheat yield changes for STSM, PDM, and RCM (bottom). The data records for each county last from 1992 to 2010. The county acronyms are: Barnim (BAR), Dahme-Spreewald (DS), Elbe-Elster (EE), Havelland (HVW), Märkisch-Oderland (MOL), Oberhavel (OHV), Oder-Spree (LOS), Ostprignitz-Ruppin (OPR), Potsdam-Mittelmark (PM), Prignitz (PR), Spree-Neiße (SPN), Teltow-Fläming (TF), and Uckermark (UM).

4.4.2 Aggregation effect

The aggregation of estimated yield changes generally increases the goodness of fit from the county to the federal state, river basin, and national scale. The extent of this improvement is shown in Tab. 2. PDMs, the models with the lowest goodness of fit measured with the NSE at the county level, gained the highest accuracy improvement by aggregation. This aggregation effect is similar for RCMs and STSMs. The aggregation to the national scale has the highest average (for all models) aggregation effect (+0.29), for river basins and federal states it is nearly the same (+0.16).

Tab. 2: Effect of spatial aggregation expressed as difference in the NSE values (ΔNSE) between the NSE at the aggregated scale (units) and the mean of N basic NSE_i across N counties: $\Delta\text{NSE} = \text{NSE}_{Unit} - N^{-1}\sum_{i=1}^{N}\text{NSE}_i$. Other terms similar to Tab. 1.

	Winter wheat			Silage maize		
Unit	STSM	PDM	RCM	STSM	PDM	RCM
Federal states (FS)						
SH	0.08	0.15	0.09	0.11	0.23	0.14
LS	0.13	0.26	0.16	0.15	0.30	0.15
NRW	0.16	0.36	0.25	0.15	0.39	0.22
HE	0.05	0.21	0.12	0.23	0.52	0.20
RP	0.08	0.22	0.10	0.11	0.32	0.17
BW	0.10	0.16	0.10	0.11	0.27	0.19
BA	0.14	0.33	0.15	0.14	0.22	0.12
SL	0.06	0.12	0.08	0.05	0.24	0.03
BB	0.09	0.22	0.12	0.08	0.14	0.08
MWP	0.13	0.29	0.17	0.12	0.22	0.14
SN	0.09	0.17	0.05	0.12	0.16	0.07
SA	0.04	0.20	0.06	0.11	0.19	0.09
TH	0.11	0.27	0.20	0.15	0.25	0.17
FSAvg	0.10	0.23	0.13	0.13	0.27	0.14
River basins (RB)						
Eider	0.07	0.16	0.13	0.04	0.12	0.04
ST	0.02	0.07	0.02	0.07	0.15	0.12
Elbe	0.18	0.43	0.17	0.20	0.38	0.14
Weser	0.16	0.29	0.20	0.21	0.45	0.22
Ems	0.12	0.23	0.16	0.05	0.20	0.03
Rhine	0.14	0.36	0.19	0.20	0.49	0.23
Maas	0.02	0.15	0.05	0.18	0.33	0.21
DAN	0.18	0.44	0.18	0.04	0.18	0.08
WP	0.11	0.28	0.12	0.12	0.23	0.11
Oder	0.01	0.06	0.05	0.02	0.06	0.02
RBAvg	0.10	0.25	0.13	0.11	0.26	0.12
National						
GER	0.19	0.44	0.21	0.27	0.49	0.16

An example of the consequences of the aggregation for the goodness of fit distributions is depicted for the winter wheat STSMs in Fig. 4. Aggregation, on the one hand, leads to a coarser resolution; on the other hand, it improves the goodness of fit for larger regions. The aggregation from county to the river basin scale shows a west-east gradient from lower to higher NSEs. The aggregation to federal states shows that the Central Uplands between Saarland and Thuringia achieve the lowest NSE values. The statistical significant STSMs are shown Fig. 4. The PDMs are all significant at $p \leq 0.05$ (F-test).

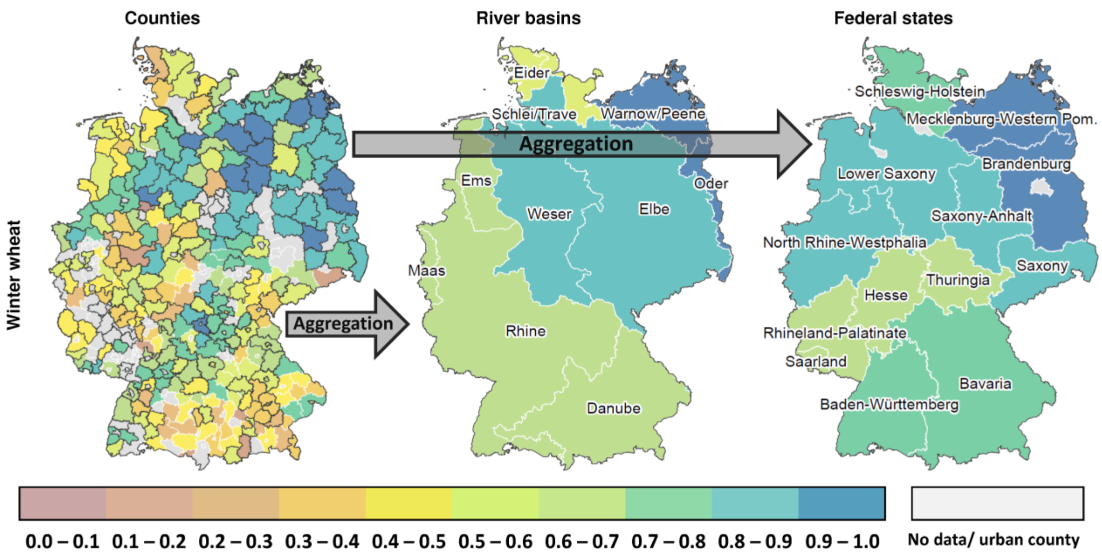


Fig. 4: Spatial distribution of the NSE for STSMs estimated for winter wheat at different scales (counties, *N*=289; federal states, *N*=13; river basins, *N*=10). Counties with significant effects (*F*-test, $p \leq 0.10$) STSMs (*N*=189) are bordered black (left).

4.4.3 Parameter heterogeneity of weather variables

The parameter distribution of all models, across scales, and for both crops is shown in Fig. 5. Due to the chosen functional form, the parameters are directly comparable. A wide range of the boxplots reflects a high spatial heterogeneity. The ranges of the parameter values are generally substantially smaller for the PDMs and the RCMs than for the STSMs. Some variables clearly show diversions from zero (*ETP* May–Jul and *SRT* May–Jul for winter wheat and *ETP* May–Jul for silage maize) while others are distributed around zero.

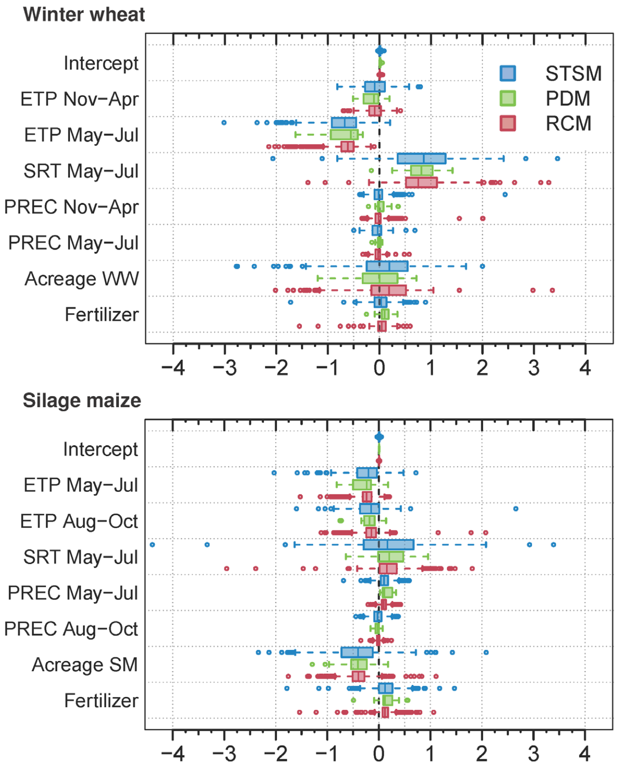


Fig. 5: Parameter distributions across counties, river basins and federal states of separately estimated time series models (STSMs), panel data models (PDMs), and random coefficient model (RCMs). The models are applied to single counties (STSMs), river basins (PDMs, RCMs) and federal states (PDMs, RCMs) for winter wheat and silage maize. The band inside the box is the median, the box represents the 25% and 75% quartile. The whiskers are defined as the maxima and the minima as long as both values are within the 1.5 interquartile-range from the median. Otherwise this range is shown and outliers outside the range are depicted as points.

For the STSMs, the spatial heterogeneous parameter variation is depicted in Fig. 6a-t. The maps of the county-individual (Fig. 6, left two columns) and the per-federal-state-averaged parameter values (Fig. 6, right two columns) often show spatial pattern, which scatter around the parameter main tendencies in Fig. 5. The larger patterns are easier to reveal following the maps with the averaged values. In particular, several parameter maps contain east-west patterns with stronger effects in eastern than in the western federal stats. That is the case for the winter wheat variables *PREC* Nov–Apr (Fig. 6 m, o), *ETP* Nov–Apr (Fig. 6 a, c), and *ETP* May–Jul (Fig. 6 e, g), and to a lesser extent for the silage maize variables *PREC* May–Jul (Fig. 6 n, p), *ETP* May-Jul (Fig. 6 b, d), and *ETP* Aug–Oct (Fig. 6 f, h). For winter wheat and silage maize, the patterns of the *SRT* May–Jul parameter reveal a north-south gradient with stronger effects in the north and weaker in the south, whereby this is more expressed for wheat than for maize.

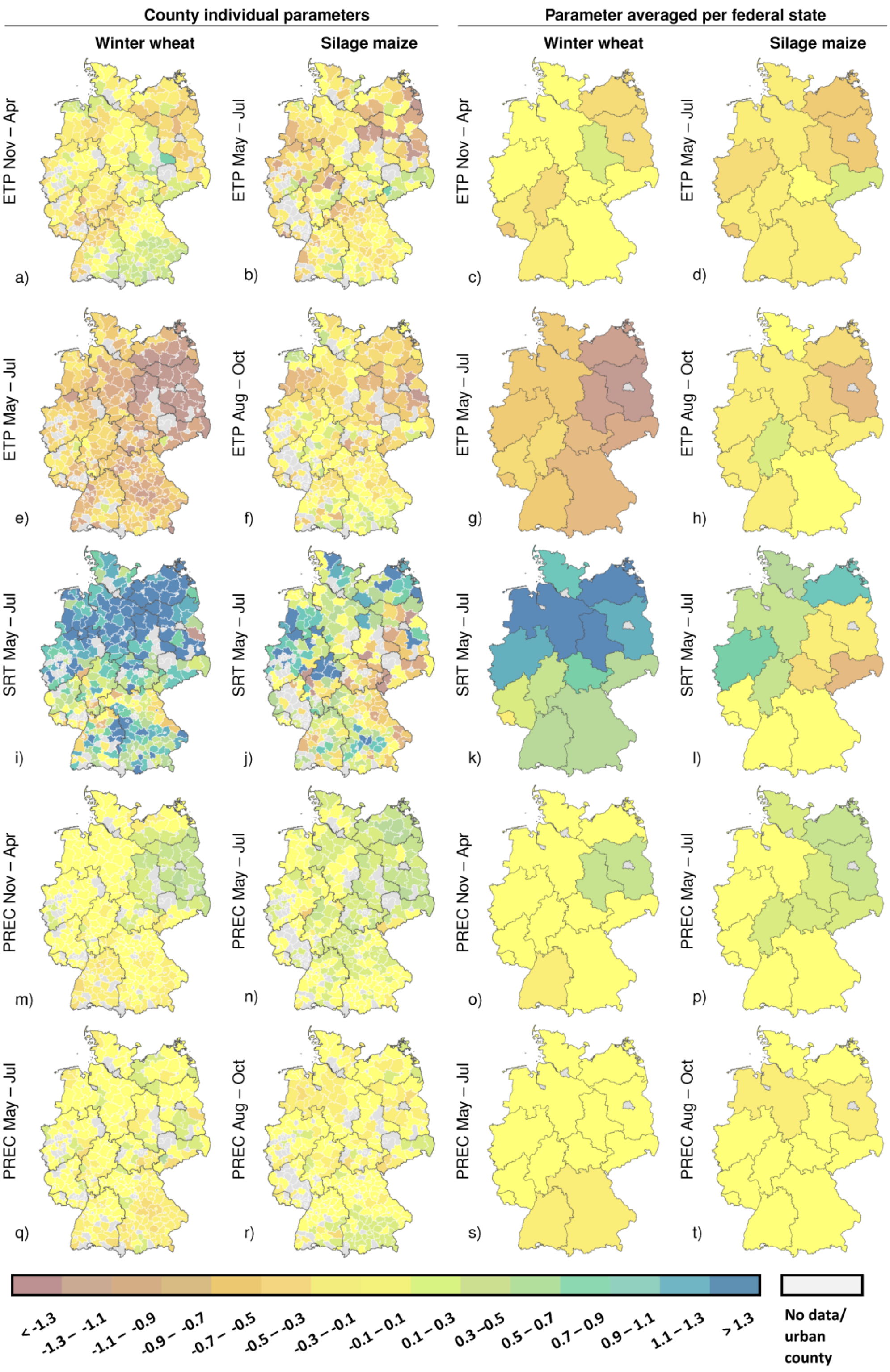


Fig. 6: Spatial distribution of the STSM parameter values for winter wheat and silage maize across counties (left two columns) and averaged per federal state (right two columns). The aggregated federal states parameter values are estimated on county scale (Eq. 2) and aggregated by arithmetic average in hindsight. The weather parameters of winter wheat are ETP Nov–Apr (a,c), ETP May–Jul (e,g), SRT May–Jul (i,k), PREC Nov–Apr (m,o), and PREC May–Jul (q,s). The weather parameters of silage maize are ETP May–Jul (b,d), ETP Aug–Oct (f,h), SRT May–Jul (j,l), PREC May–Jul (n,p), and PREC Aug–Oct (r,t).

4.4.4 Statistical tests

The validity of a spatially distributed modeling approach is statistically confirmed by the LM-test. The county yields do not dependent on each other. The Cobb–Douglas function is mostly not error-specified as the functional form (RESET) when testing the STSMs of both crops (SI Fig. S.1). This means quadratic variables would not improve the model goodness of fit. According to the RESET, the winter wheat PDMs are partly error-specified and the silage maize PDMs are often error-specified (SI

Tab. S.1). Both autocorrelation and heteroskedasticity do not occur in the majority of the STSMs. Their residuals are mostly normally distributed (SI Fig. S.1). For the PDMs, autocorrelation and/ or heteroskedasticity are common (SI Tab. S.1). The quality of the RCMs can be assessed using the NSEs in Tab. 1.

Several highly significant correlations exist among the transformed, weather and economic variables. For instance, *SRT* is moderately correlated with *ETP* and precipitation (0.67^{***}, -0.52^{***}), as well as precipitation and *ETP* (-0.58^{***}), in the period May to Jul. Furthermore, fertilizer price and acreage of winter wheat and silage maize are strongly correlated (0.72^{***}, 0.87^{***}). However, a test of multicollinearity (condition index) resulted in values always lower than 12 (lower 8 in 97.5% of all cases). Following Belsley et al. (1980), values beyond 30 are an indicator for multicollinearity. Further information about the results of the statistical tests and the correlation of all exogenous variables are presented in the SI Fig. S.1, Tab. S.1 and S.2.

4.5 Discussion

4.5.1 Goodness of fit and yield variability between crops and regions

We investigate and validate three statistical models (STSM, PDM, and RCM) according their robustness for short and medium-term yield assessments. These three models are tested to capture the weather -related yield variability of winter wheat and silage maize in Germany. All models are able to satisfactorily reproduce the temporal and spatial variability of yields. In general, the differences in goodness of fit between winter wheat and silage maize are low. The models of regions with higher yield variability have generally scored the highest NSE. Thus, a clear west–east NSE gradient for the federal state and river basin scale is observable. We found a very strong positive correlation between the goodness of fit and yield variability for silage maize (0.86) and a strong positive correlation for winter wheat (0.66). This relationship is visible across the federal states, but not between the two crops.

The STSMs, which are the simplest models, perform best, followed by RCMs and PDMs. The ranking holds true for estimations, validations, and expanded validations, which indicates that STSMs are robust to missing data despite their higher parameter numbers. The advantage in robustness of STSMs compared to the other two approaches originates in the estimation method. The STSMs are separately estimated for each county. The parameters of STSMs vary more across the counties than those of RCMs and PDMs. Thus, STSMs can reproduce better extreme county individual yield anomalies than RCMs and PDMs, which might be beneficial for the adequate reproduction of heterogeneous spatial conditions (Beck and Katz, 2007; Butler and Huybers, 2013). In contrast, the PDM parameters are estimated for the entire dataset to reproduce the full range of yield variability. Their parameters reproduce rather the mean yield level than the full range yields variability. Nevertheless, the higher parameter number and the lower degree of freedom of the STSMs are a potential source of parameter instability when the time series becomes shorter. Under such conditions, RCMs and PDMs might behave more robustly than STSMs (Reidsma et al., 2007; You et al., 2009). Furthermore, projections at the county scale beyond the observed yield variability might be more biased by single regional events at that scale when using STSMs compared to PDMs and RCMs. Such a disadvantage might lose relevance with increasing spatial aggregation of the STSM results.

4.5.2 Aggregation effect

A substantially higher goodness of fit is achieved by all models after aggregation from county to larger spatial units. The winter wheat STSMs (+0.10) show the smallest aggregation effect in comparison to the respective RCMs (+0.13) and to the PDMs (+0.23) at the averaged federal state scale. For silage maize and the other scales, the effects are similar. In tendency, the advantages of STSMs at the lowermost scale (here county) lose relevance at the more aggregated scale. Thus, aggregation has a slighter effect on the goodness of fit for STSMs than of RCMs and PDMs. Woodard and Garcia (2008), Lobell and Burke (2010), and Hanus (1978) have noted the aggregation effect before. Conradt et al. (2015) used the parameter vectors of our STSMs for cluster analyses to define optimized PDM aggregations independently from federal states or river basin scales. This could again, but only slightly, add to the overall goodness of fit; at least the county-specific fidelity of the estimations became much more homogeneous. In our approach, only the estimated outcomes are spatially aggregated and not the exogenous variables. Aggregated exogenous variables can lead to an underestimation of the weather effect (Garcia et al., 1987), to decreased variability, and erroneous results (Finger, 2012).

4.5.3 Parameter distributions and patterns

Winter wheat is more responsive than silage maize to higher evaporative demand during spring and summer as indicated by the more negative values for *ETP* May–Jul. That might be due to the more developed plant canopy. After closing the canopy (Aug–Oct), the silage maize shows a clearer negative impact of higher *ETP*. For the *ETP* related vapor pressure deficit, a negative yield impact is also shown by Lobell et al. (2014) and Roberts et al. (2012). Consistent with this explanation, the less developed silage maize in May to Jul (early vegetative development, between emergence and canopy closure) is more sensitive to lower water supply than winter wheat during that time (*PREC* May–Jul). For winter wheat, a similar effect is observable during the early plant development stages, in particular in the eastern parts of Germany. This region is marked by sandy soils with low water holding capacity and low precipitation levels. These conditions lead to a higher sensitivity of crop yields (high yield variability) to inter-annual changes of water supply. Wessolek and Asseng (2006) also show the importance of this limited water supply for winter wheat in north-east Germany. The importance of the water supply for winter wheat and silage maize in Germany is also emphasized by Kersebaum and Nendel (2014) and Wolf and Diepen (1994).

Furthermore, winter wheat benefits more than silage maize from higher *SRT* May–Jul values during that period. This might reflect the higher temperature sensitivity of light respiration of C_3- (e.g. wheat) than C_4-crops (e.g. maize). As a consequence, lower temperatures at the same radiation levels and higher radiation levels at the same temperature levels (increasing *SRT*) function more positively on winter wheat than on silage maize (Long et al., 2006; Rötter and van de Geijn, 1999). However, Conradt et al. 2015 could increase the model performance by decoupling radiation and temperature back into two model parameters considering regional exceptions to these general patterns. A detailed discussion of the spatial parameter patterns for winter wheat and silage maize is in the SI S.6.

Surprisingly, the parameters for precipitation indicate a small yield effect compared to the other factors considered for both crops. A possible explanation is offered by the variability differences among variables. In our dataset, the transformed precipitation from May to Jul varies by ±43% (relative SD), while *ETP* and *SRT* only vary by ±17% and ±8%, respectively. Due to their high relative SD, small parameter values are estimated for precipitation. However, for the assessment of weather-yield impacts the explained yield variability ($\beta_j \ \log x'_j$) is more important than the parameter size (β_j). A further analysis shows that the yield variability explained by precipitation is substantially larger in comparison to the other variables (SI Fig. S.2). The result possibly explains the small yield impact for precipitation in Europe reported by Moore and Lobell (2014). They have drawn their conclusion solely from the parameter size, but not from the explained variability.

Generally, the STSM parameters show parameter patterns on a broader scale, but also county specific heterogeneity. The spatial parameter patterns can be explained by linear relationships between yield and exogenous variables, because of spatially heterogeneous levels of the exogenous variables. The county parameters do not deviate ideally from the average parameters of the broader patterns. In our case, the parameters may also reflect individual factor influences, which are not considered in the model. These influences are collinear with the considered variables but not relevant in the majority of the counties (county individual time variant effects). The impact of those factors may lead to spatial heterogeneity between neighboring counties that cannot be explained by differences in soil characteristics or cropping structure. For instance, the possible collinear influences might be catch crops (*ETP*), weeds, pests, and diseases (*SRT*), or irrigation (*PREC*).

4.5.4 Model application in climate impact studies

Our modeling scheme allows a direct interpretation of the spatial parameter variability and a usage for crop yield assessments with seasonal- and medium-term climate projections. Both characteristics are based on a consequent usage of changes instead of absolute values, which contributes to the methodological novelty of the approach. The parameter values and patterns can be used to prove the plausibility of the model outcome. The feasibility of plausibility test is supported by a variable definition that reflects major climate impacts on potential growth and stress related to limited water supply. The selected variables might be meaningful also in other wheat and maize growing regions. However, an adjustment of the temporal division to the regional crop calendar is necessary. The use of changes makes the model also insensitive to systematic errors in data from climate simulations. This insensitivity does not avoid flawed yield projections of flawed climate simulations. However, considering the necessary effort of bias correction and the often nontransparent procedure (Lobell, 2013), our models are an option for using the outcome of climate simulations in advance of a later bias correction. It is not the solution for the bias problem of climate simulations, but an improvement for their technical handling.

Butler and Huybers (2013) show that the impact of temperature on US maize yields is very sensitive in respect to the latitude and the regional climate conditions. STSMs, PDMs and RCMs should be principally applicable for such conditions in order to project climate impacts. However, the advantages and limitations of each model should be kept in mind. Our approach implicitly accounts for the different yield sensitivities of vegetative and reproductive growth periods to climate changes. Any further detailed resolution of the phenological development might be beneficial in statistical analysis (Butler and Huybers, 2015). However, yield projection of statistical yield models would require phenological development data also for the future. Since phenological models (Ma et al., 2012) and climate simulations (Lobell, 2013) are becoming robust only at broader temporal and spatial resolution, we use monthly averaged phenological dates, to make our models suitable for future projections.

Our statistical models project future yield changes on the basis of the current system. Several factors and factor relationships that are unknown today might play a major role in the future and are not included in the model. In our model set-up, we focused on the representation of regularly returning yield impacts of climate variables that can be reliably received from climate models. The impact of extreme weather events that affected the crop yield only episodically in the past but will become regular disturbances in the future might be underestimated. Furthermore, if climatic change passes thresholds, crop yields might be seemingly insensitive due to unconsidered climate impacts during the parameter estimation of our crop yield models (Blanc and Sultan, 2015; Rötter et al., 2011). Yield effects of technological change and the impact of higher CO_2 (by stimulating crop growth and increasing water use efficiency) are also not included in the model. They could be taken into account by introducing a post-processing to the model output using external correction factors as exercised by Wechsung et al. (2008).

4.6 Conclusion

Our suggested approach can be used for seasonal yield forecasts and climate impact projections on crop yields. For short and medium term climate assessments, we investigate and validate three types of statistical crop yield models (STSM, PDM, and RCM). These models are suitable for a combination with biased climate simulations and avoid explicit modeling of crop yield trends. Our approach is thoroughly based on relative changes of yields and yield influencing factors. Our models can reproduce past regional yield variability; they are robust to data fragmentation and show reasonable parameter patterns at aggregated scales. Although STSMs have shown the best performance at the aggregated scale, the model assessments at the county scale should only be used as technical intermediate steps but not as projections. The suggested regression models might be applicable to calculate weather-related yield risks and thus support investment decisions (e.g. in irrigation systems) and risk pricing (e.g. of harvests or agricultural commodity futures for farmers, traders, or insurance companies).

4.7 References

Angulo, C. et al., 2013. Implication of crop model calibration strategies for assessing regional impacts of climate change in Europe. Agricultural and Forest Meteorology, 170: 32-46.

Asseng, S. et al., 2013. Uncertainty in simulating wheat yields under climate change. Nature Climate Change, 3: 827–832.

Bakker, M.M., Govers, G., Ewert, F., Rounsevell, M. and Jones, R., 2005. Variability in regional wheat yields as a function of climate, soil and economic variables: Assessing the risk of confounding. Agriculture, Ecosystems & Environment, 110(3-4): 195-209.

Beck, N. and Katz, J.N., 2007. Random Coefficient Models for Time-Series–Cross-Section Data: Monte Carlo Experiments. Political Analysis, 15: 182-195.

Belsley, D.A., Kuh, E. and Welsch, R.E., 1980. Regression diagnostics: identifying influential data and sources of collinearity. Wiley, New York, 300 p.

Blanc, E. and Sultan, B., 2015. Emulating maize yields from global gridded crop models using statistical estimates. Agricultural and Forest Meteorology, 214–215: 134-147.

Butler, E.E. and Huybers, P., 2013. Adaptation of US maize to temperature variations. Nature Climate Change, 3: 68-72.

Butler, E.E. and Huybers, P., 2015. Variations in the sensitivity of US maize yield to extreme temperatures by region and growth phase. Environmental Research Letters, 10: 1-8.

Castellvi, F., Perez, P.J., Stockle, C.O. and Ibaiiez, M., 1997. Methods for estimating vapor pressure deficit at a regional scale depending on data availability. Agricultural and Forest Meteorology, 87(4): 243-252.

Castellvi, F., Perez, P.J., Villar, J.M. and Rosell, J.I., 1996. Analysis of methods for estimating vapor pressure deficits and relative humidity. Agricultural and Forest Meteorology, 82: 29-45.

Chipanshi, A. et al., 2015. Evaluation of the Integrated Canadian Crop Yield Forecaster (ICCYF) model for in-season prediction of crop yield across the Canadian agricultural landscape. Agricultural and Forest Meteorology, 206: 137-150.

Chmielewski, F. and Köhn, W., 2000. Impact of weather on yield components of winter rye over 30 years. Agricultural and Forest Meteorology, 102: 253-261.

Conradt, T., Gornott, C. and Wechsung, F., 2015. Extending and improving regionalized winter wheat and silage maize yield regression models for Germany: enhancing the predictive skill by panel definition through cluster analysis. Agricultural and Forest Meteorology, (216) 68–81.

Croissant, Y. and Millo, G., 2008. Panel Data Econometrics in R: The plm Package. Journal of Statistical Software, 27(2): 1-43.

Dielman, T.E., 1983. Pooled Cross-Sectional and Time Series Data: A Survey of Current Statistical Methodology. The American Statistician, 37(2): 111-122.

Dixon, B.L., Hollinger, S.E., Garcia, P. and Tirupattur, V., 1994. Estimating Corn Yield Response Models to Predict Impacts of Climate Change. Journal of Agricultural and Resource Economics, 19(1): 58-68.

DWD, 2011. Daily weather data, 1951 - 2010. German Weather Service.

DWD, 2015. Aktueller Stand der Phänologie in Deutschland, 1992 - 2013. German Weather Service, Retrieved from: http://www.dwd.de/phaenologie (11.06.2015).

Estes, L.D. et al., 2013. Projected climate impacts to South African maize and wheat production in 2055: a comparison of empirical and mechanistic modeling approaches. Global Change Biology, 19(12): 3762–3774.

Federal Statistical Office, 2013. Ackerland nach Hauptfruchtgruppen und Fruchtarten.

Finger, R., 2012. Biases in Farm-Level Yield Risk Analysis due to Data Aggregation. German Journal of Agricultural Economics, 61: 30-41.

Garcia, P., Offutt, S.E., Pinar, M. and Changnon, S.A., 1987. Corn Yield Behavior - Effects of Technological Advance and Weather-Conditions. Journal of climate an applied meteorology, 26: 1092-1102.

Gornott, C. and Wechsung, F., 2015. Niveauneutrale Modellierung der Ertragsvolatilität von Winterweizen und Silomais auf mehreren räumlichen Ebenen in Deutschland. Journal für Kulturpflanzen, 67: 205-223.

Hanus, H., 1978. Vorhersage von Ernteerträgen aus Witterungsdaten in den Ländern der EG. eurostat, 21: i-54.

Haude, W., 1955. Zur Bestimmung der Verdunstung auf möglichst einfache Weise. Mitteilungen des Deutschen Wetterdienstes 11.

Iizumi, T. et al., 2013. Prediction of seasonal climate-induced variations in global food production. Nature Climate Change, 3: 904–908.

Kaufmann, R.K. and Snell, S.E., 1997. A Biophysical Model of Corn Yield: Integrating Climatic and Social Determinants. American Journal of Agricultural Economics, 79(1): 178-190.

Kersebaum, K.C. and Nendel, C., 2014. Site-specific impacts of climate change on wheat production across regions of Germany using different CO2 response functions. European Journal of Agronomy, 52: 22-32.

Krause, J., 2008. A Bayesian approach to German agricultural yield expectations. Agricultural Finance Review, 68: 9-23.

Krause, P., Boyle, D.P. and Bäse, F., 2005. Comparison of different efficiency criteria for hydrological model assessment. Advances in Geosciences, 5: 89-97.

Lee, B.-H., Kenkel, P. and Brorsen, B.W., 2013. Pre-harvest forecasting of county wheat yield and wheat quality using weather information. Agricultural and Forest Meteorology, 168: 26-35.
Lobell, D.B., 2007. Changes in diurnal temperature range and national cereal yields. Agricultural and Forest Meteorology, 145: 229–238.
Lobell, D.B., 2013. Errors in climate datasets and their effects on statistical crop models. Agricultural and Forest Meteorology, 170: 58-66.
Lobell, D.B. and Asner, G.P., 2003. Climate and Management Contributions to Recent Trends in U.S. Agricultural Yields. Science, 299(5609): 1032.
Lobell, D.B. and Burke, M.B., 2010. On the use of statistical models to predict crop yield responses to climate change. Agricultural and Forest Meteorology, 150(11): 1443-1452.
Lobell, D.B. et al., 2013. The critical role of extreme heat for maize production in the United States. Nature Climate Change, 3(5): 497-501.
Lobell, D.B. et al., 2014. Greater Sensitivity to Drought Accompanies Maize Yield Increase in the U.S. Midwest. Science, 344(6183): 516-519.
Long, S.P., Zhu, X.G., Naidu, S.L. and Ort, D.R., 2006. Can improvement in photosynthesis increase crop yields? . Plant Cell and Environment, 29: 315-330.
Ma, S., Churkina, G. and Trusilova, K., 2012. Investigating the impact of climate change on crop phenological events in Europe with a phenology model. International Journal of Biometeorology, 56(4): 749-763.
Monteith, J.L., 1977. Climate and the efficiency of crop production in Britain. Philosophical Transactions of the Royal Society of London, 281: 277-294.
Moore, F.C. and Lobell, D.B., 2014. Adaptation potential of European agriculture in response to climate change. Nature Climate Change, 4: 610–614.
Mueller, N.D. et al., 2012. Closing yield gaps through nutrient and water management. Nature, 490: 254-257.
Oury, B., 1965. Allowing for Weather in Crop Production Model Building. Journal of Farm Economics, 47(2): 270-283.
Palosuo, T. et al., 2011. Simulation of winter wheat yield and its variability in different climates of Europe: A comparison of eight crop growth models. European Journal of Agronomy, 35(3): 103-114.
Ray, D.K., Gerber, J.S., MacDonald, G.K. and West, P.C., 2015. Climate variation explains a third of global crop yield variability. Nature Communications, 6: 1-9.
Reidsma, P., Ewert, F. and Lansink, A.O., 2007. Analysis of farm performance in Europe under different climatic and management conditions to improve understanding of adaptive capacity. Climatic Change, 84(3-4): 403-422.
Roberts, M.J., Schlenker, W. and Eyer, J., 2012. Agronomic weather measures in econometric models of crop yield with implications for climate change. American Journal of Agricultural Economics, 95(2): 236-243.
Rötter, R. and van de Geijn, S.C., 1999. Climate Change Effects on Plant Growth, Crop Yield and Livestock. Climatic Change, 43: 651-681.
Rötter, R.P., Carter, T.R., Olesen, J.E. and Porter, J.R., 2011. Crop-climate models need an overhaul. Nature Climate Change, 1: 175–177.
Schlenker, W. and Roberts, M.J., 2009. Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. Proceedings of the National Academy of Sciences of the United States of America, 106(37): 15594-15598.
Schrödter, H., 1985. Verdunstung: Anwendungsorientierte Meßverfahren und Bestimmungsmethoden. Springer, 204 p.
Sonntag, D., 1990. Important new Values of the Physical Constants of 1986, Vapour Pressure Formulations based on ITS-90, and Psychrometer Formulae. Meteorologische Zeitschrift, 4(5): 340-344.
Statistical Offices of the Federation and the Länder, 2013a. Datensatz Anbaufläche (Feldfrüchte und Grünland): Deutschland, Jahre, Fruchtarten 1991-2007.
Statistical Offices of the Federation and the Länder, 2013b. Hektarerträge ausgewählter landwirtschaftlicher Feldfrüchte - Jahressumme - regionale Tiefe: Kreise und krfr. Städte.
Statistical Offices of the Federation and the Länder, 2013c. Einkaufspreise landwirtschaftlicher Betriebsmittel.
Wechsung, F., Lüttger, A.B. and Hattermann, F., 2008. Projektionen zur klimabedingten Änderung der Erträge von einjährigen Sommer-und Winterkulturen des Ackerlandes am Beispiel von Silomais und Winter-weizen, Die Ertragsfähigkeit ostdeutscher Ackerflächen unter Klimawandel. PIK-Report 112, p. 18-32.
Wessolek, G. and Asseng, S., 2006. Trade-off between wheat yield and drainage under current and climate change conditions in northeast Germany. European Journal of Agronomy, 24(4): 333-342.
Wolf, J. and Diepen, C.A.v., 1994. Effects of climate change on silage maize production potential in the European Community. Agricultural and Forest Meteorology, 71: 33-60.
Woodard, J.D. and Garcia, P., 2008. Weather Derivatives, Spatial Aggregation, and Systemic Risk: Implications for Reinsurance Hedging. Journal of Agricultural and Resource Economics, 33(1): 34-51.
Wooldridge, J.M., 2013. Introductory econometrics. a modern approach. South Western Cengage Learning, 868p.
You, L., Rosegrant, M.W., Wood, S. and Sun, D., 2009. Impact of growing season temperature on wheat productivity in China. Agricultural and Forest Meteorology, 149(6-7): 1009-1014.

5 Global evaluation of a semi-empirical model for yield anomalies and application to within-season yield forecasting

(Semi-empirical modeling of yield anomalies)

Bernhard Schauberger[1,2*], Christoph Gornott[1], Frank Wechsung[1]

[1] Potsdam Institute for Climate Impact Research (PIK)

[2] Laboratoire des Sciences du Climat et de l'Environnement, Institut Pierre-Simon Laplace (IPSL)

[*] Corresponding author

5.1 Abstract

Quantifying the influence of weather on yield variability is decisive for agricultural management under current and future climate anomalies. We extended an existing semi-empirical modeling scheme that allows for such quantification. Yield anomalies, measured as inter-annual differences, were modeled for maize, soybeans and wheat in the US and 32 other main producer countries. We used two yield data sets, one derived from reported yields and the other from a global yield data set deduced from remote sensing. We assessed the capacity of the model to forecast yields within the growing season. In the US, our model can explain at least two thirds (63-81%) of observed yield anomalies. Its out-of-sample performance (34-55%) suggests a robust yield projection capacity when applied to unknown weather. Out-of-sample performance is lower when using remote-sensing derived yield data. The share of weather-driven yield fluctuation varies spatially, and estimated coefficients agree with expectations. Globally, the explained variance in yield anomalies based on the remote-sensing data set is similar to the US (71-84%). But the out-of-sample performance is lower (15-42%). The performance discrepancy is likely due to shortcomings of the remote-sensing yield data since it diminishes when using reported yield anomalies instead. Our model allows for robust forecasting of yields up to two months before harvest for several main producer countries. An additional experiment suggests moderate yield losses under mean warming, assuming no major changes in temperature extremes. We conclude that our model can detect weather influences on yield anomalies and project yields with unknown weather. It requires only monthly input data and has a low computational demand. Its within-season yield forecasting capacity provides a basis for practical applications like local adaptation planning. Our study underlines high-quality yield monitoring and statistics as critical prerequisites to guide adaptation under climate change.

Keywords: yield anomaly, maize, wheat, soybeans, global, weather, semi-empirical model, forecast

5.2 Introduction

Strongly varying crop yields can endanger farmers' livelihoods and can lead to national production shortages. Yields are determined by weather and agronomic management influences as well as by stress factors like pests or diseases. For calculating crop yields under current or a changing climate it is important to quantify these influences. Therefore we devise a semi-empirical modeling scheme which allows for quantifying weather influences with high explained variance. We use two different yield data sets with different qualities, one based on reported yield data and the other on remote sensing combined with yield statistics. We show the ability of the model to predict yield anomalies up to two months before harvest.

Two approaches are widely used to simulate crop yields (Di Paola et al., 2016, Jones et al., 2016, Lobell and Burke, 2010). Process-based models simulate physiological processes like carbon assimilation to calculate yields. Statistical models correlate yields with yield-determining factors to elicit contributions of individual factors. Both approaches, and hybrids between them, can aid in understanding and forecasting weather-related yield variability (Liu et al., 2016). Their application to conditions (e.g. climate) out of the training scope is a contested area, however (Lobell and Burke, 2010, Rötter et al., 2011).

Here we extend an existing statistical framework for modeling inter-annual yield variability. The approach is "semi"-empirical as known physiological influences are reflected in the exogenous variables, following the naming of Rahmstorf (2007). The concept was introduced in Wechsung et al. (2008) and later successfully applied to German maize and winter wheat yields (Gornott and Wechsung, 2016). We extend the model by adding temperature-stress related variables, using more crops, applying it to 34 countries and providing two application cases: forecasting yield anomalies up to two months before harvest and gauging of yield losses under moderately increased temperatures.

We analyze four staple crops: maize, wheat (spring and winter separately) and soybeans, which cover approx. 34% of the global harvested area (Portmann et al., 2010). We use reported crop yield data in seven countries and a global gridded yield data set that downscaled reported yield statistics utilizing satellite data (here used for 33 countries). Subnational yield data are needed for quantifying spatial differences of yield influences. Though these data are increasingly available, there are still data-scarce regions especially in developing countries. The global and publicly available data set supplied by Iizumi et al. (2013b) might serve as alternative. The dataset uses annual remote sensing information to downscale national and subnational yield statistics. The algorithms utilized therein to separate reflectance data spatially and temporally into crops or vegetation necessarily introduce uncertainty, which increases with the share of other vegetation types in grid cells. Despite these caveats we test the potential of this global gridded data set for quantifying yield anomalies, as it may be helpful when subnational yield data are not accessible.

We apply a two-step procedure: the model performance is first analyzed in depth in the US and then, second, extended to all main producing nations. We start with US yields, since the high-quality yield data base curated by the US Department of Agriculture (USDA, 2015) allows for rigorous model evaluation. The model is applied in parallel to the USDA and the Iizumi et al. (2013b) data. The US are one of the largest crop producers (FAO, 2016) and have highly diverse climate and soils. We employ one model specification based on selection results by Gornott and Wechsung (2016), but test its sensitivity regarding variations in yield-influencing factors and transformation of variables. Additionally, we include penalty terms for heat and frost.

Instead of absolute yields we consider yield anomalies to remove trends, systematic biases and time-invariant farm- or county-specific influencing factors. Normalizing anomalies of yield and exogenous variables by the logarithm allows a comparison of influences across scales and variables. Only weather variables are included in the model, explicitly neglecting agronomic influences like acreage, shifting land use or fertilizer application on inter-annual yield fluctuation (Mueller et al., 2012, Ray et al., 2015). But these data do not increase model performance in Germany (Conradt et al., 2016) and are difficult to obtain as time series on a spatially explicit level with large spatial coverage; they would therefore enlarge uncertainty. We only use monthly weather values which are deemed to provide more reliable information than daily weather data from models due to aggregation effects (Kilsby et al., 2007, Lobell, 2013, Maurer et al., 2010). This also avoids the use of downscaling methods when using climate model outputs (Glotter et al., 2014, Iizumi et al., 2012).

5.3 Materials and methods

5.3.1 Input data

Yield data

We employed two sets of yield data for maize, soybeans, spring and winter wheat (all in t/ha). For the US we used either USDA (USDA, 2015) yields at county level, from 1980 to 2010, or gridded yield data from Iizumi et al. (2013b) from 1982 to 2006 (henceforth “GGYD” for “Gridded Global Yield Data”). Both were re-gridded to 0.5° spatial resolution (about 50 km at the equator) to match with the resolution of the weather and land-use data. USDA county yields were assigned to each 0.5° grid cell that completely fall within a county or intersect with its boundaries; yields for grid cells intersecting with several counties were averaged. GGYD yields are provided at 1.125° resolution and were interpolated to 0.5° with second order conservative remapping (preserving fluxes and spatial gradients). Additional county-level yields for Germany, Russia, Tanzania, Australia, Brazil and Burkina Faso (from the respective statistical offices) allowed for further model and yield data quality assessments. National yield time series from FAO (FAO, 2016) were used for comparison of aggregated yield time series.

We considered those countries as main producers (Fig. 1, SI Tab. S3) which, sorted by total production, together accounted for more than 90% of world production for a specific crop between 2000 and 2011 (FAO, 2016).

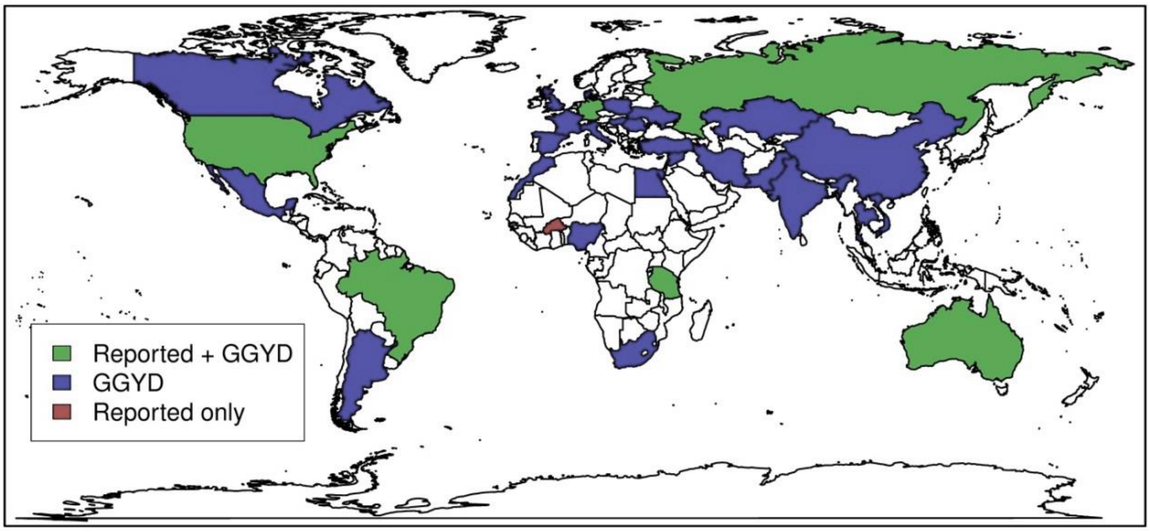


Fig. 1: World map of countries analyzed in this study. Colors of countries denote whether GGYD and reported yields (green), only GGYD yields (blue) or only reported yields (red) are used in this study. Countries in white are no main producers and not analyzed.

Weather data

We used AgMERRA climate data (Ruane et al., 2015) at 0.5° spatial and monthly temporal resolution, providing minimum, maximum and average temperature, precipitation and shortwave radiation from 1980 to 2010. AgMERRA has been designed for use in agricultural research focusing on reproducing both average and extreme values.

Growing season and land-use data

We utilized static MIRCA2000 crop- and irrigation-specific land-use fractions around 2000 on 0.5° spatial resolution (Portmann et al., 2010). Growing seasons were also taken from MIRCA2000, using the sub-crop with the largest harvested area. Winter and spring wheat were distinguished by their growing season length: eight or more months were classified as winter wheat, four months or less as spring wheat. Remaining ambiguities were resolved by considering the sub-crop with the maximum (minimum) growing season length as winter (spring) wheat. Soybeans have a prolonged flowering period (Ritchie et al., 1993) at the transition between vegetative and reproductive season. Although it could be physiologically reasonable, we restrained from reflecting this period in a separate set of exogenous variables to avoid collinearities and rank deficiencies (many variables for few data).

5.3.2 Regression scheme

Definition

We applied an ordinary least squares (OLS) regression scheme based on the Cobb-Douglas production function with different model specifications. The function relates inter-annual changes of crop yields to a product of inter-annual changes of weather variables (equation 1; SI equation SE3). The natural logarithm linearizes all terms into a sum.

$$\log y_t{}' = \log \beta_0 + \textstyle\sum_{j=1}^{J} \beta_j \log x'_{jt} + \log u'_t \text{, with } j = 1, \dots, J \text{ and } t = 1, \dots, M \tag{1}$$

Variables are yield (y), weather (x_j) and error term (u). Estimated coefficients are $\beta_{0\dots J}$ and denote intercept (β_0) and weather influences. All variables are provided per grid cell. Years are indexed with t. Anomalies are denoted with a prime ($'$). We calculated yield anomalies as first differences ($y'_t = y_t - y_{t-1}$) between adjacent years, making an explicit time variable obsolete. We used two regression methods: STSM (Separate Time Series Model) and PDM (Panel Data Model). While STSM estimates an independent model for each grid cell, the PDM parametrizes relationships across grid cells, allowing for spatial variation in mean yields with grid cell-specific fixed effects. These choices are justified by earlier results (Conradt et al., 2016, Gornott and Wechsung, 2016) and the similarity of results under different techniques (SI Section 3). Whether spatial correlation poses a problem for the PDM method is tested (see below). In the US we considered nine climatic regions (SI Fig. s S1-2). Other, larger main producers were split into administrative boundaries for PDM estimation; for all others only one national PDM was estimated (SI Tab. S3).

Exogenous variables

Exogenous variables either describe potential growth or stress factors that reduce growth, included for their known physiological relevance. They are tested for statistical significance, but the model formulation stays constant. We therefore consider the model as “semi”-empirical following the argumentation of Rahmstorf (2007). A combined temperature-radiation variable relates yields to potential growth. Temperature-normalized solar radiation (SRT, equation 2) is used to account for co-linearity in both variables. Killing (KDD) and freezing degree days (FDD) were added to better account for the non-linear influence of extreme temperatures on crop yields (Barlow et al., 2015, Schlenker and Roberts, 2009). They are defined as the temperature sum above or below a crop-specific threshold, respectively (equations 3,4). The KDD threshold T^{KDD} was 32°C for all crops, while the FDD threshold T^{FDD} was -15°C for the two wheat types and 0°C for maize and soybeans (Hatfield et al., 2011, Luo, 2011, Porter and Gawith, 1999, Sanchez et al., 2014).

$$SRT = \frac{R_S}{T_{avg} + 20} \tag{2}$$

$$KDD = \sum_{d=1}^{N} \max(T_d - T^{KDD}; 0) \tag{3}$$

$$FDD = \sum_{d=1}^{N} \min(T_d - T^{FDD}; 0) \tag{4}$$

Further stress variables comprised potential evapotranspiration (PET) and precipitation. Both variables map the yield-reducing effect of inadequate demand and supply of water by PET and precipitation, respectively. PET was calculated from VPD according to Haude (1955) as in Gornott and Wechsung (2016) except that the month-specific correction factor f_H was considered constant for the sake of a simpler model. For winter wheat only the reproductive part of SRT was considered, while for the other crops only the vegetative part was used. The full regression specification is provided in SI section 2. Further agronomic justifications are provided in Gornott and Wechsung (2016). Economic variables like fertilizer price and harvested area were not considered since these only added little explanatory power in Germany (Conradt et al., 2016) and are generally not available on larger areas across the world.

PET and precipitation were split between the vegetative and reproductive part of the growing season. The identification of both parts was based on phenological heat units. The first month of the reproductive period was defined as the first month where the temperature sum, accumulated over the growing season until this month, exceeds 50% of the total temperature sum, accumulated over the whole growing season (supplementary equations SE4,5).

Aggregation

After estimation yield anomaly time series (observed, predicted and one-out-of-sample predicted yield anomalies) were aggregated from grid cells to climate regions or countries (supplementary equations SE1,2). Aggregation was performed unweighted, i.e. treating each grid cell as equal, or weighted by land-use patterns according to MIRCA2000. Performance measures (see below) were then calculated for aggregated time series.

5.3.3 Model evaluation

Performance

Six performance indicators were calculated: coefficient of determination (R^2), root mean square error (RMSE), Nash-Sutcliffe efficiency (NSE), one-out-of-sample R^2 (henceforth: R^2_{OI}), out-of-temperature R^2 (R^2_{OOT}) and out-of-precipitation R^2 (R^2_{OOP}). The first three are standard model evaluation indices and measure the explained variance, the mean deviation and a combined measure of model bias and variability, respectively. They indicate the capacity of the model to explain yield anomalies, which is important for interpreting coefficients. R^2_{OI} was calculated by subsequently and separately stripping each year from the estimation data, estimating the model with the reduced data and eventually predicting yield anomalies for the stripped year with this reduced model. R^2_{O1} thus indicates the model's capacity to project yields from weather data that have not been used for model training. R^2_{OOT} and R^2_{OOP} were similarly calculated by omitting the six first-differences towards and from the three warmest (driest) years, defined by highest growing season mean temperature (lowest precipitation over PET). Thus the model was trained on six yield anomalies less and was then used to predict these missing anomalies. The correlation between these predicted and observed anomalies in only the warmest (driest) years, calculated across aggregation regions, indicates the capacity to project yield anomalies under warmer (drier) climate. Performance measures were calculated on nationally aggregated time series, but are also available for each grid cell.

Statistical tests

The adequacy of the linear model for capturing yield anomalies was examined with six statistical tests. The regression equation specification error test (RESET) evaluated whether quadratic variables would improve the model. The Lagrange multiplier test according to Breusch–Pagan (LM) was used to examine spatial independence of the data. The Breusch–Godfrey test was applied to assess autocorrelation and the Breusch–Pagan test to probe heteroscedasticity (Croissant and Millo, 2008, Wooldridge, 2013). Normal distribution of residuals was tested using the Shapiro–Wilk test. Whether multicollinearity of exogenous variables poses a problem was assessed with the condition index following Belsley et al. (1980). All analyses were performed with R (R Core Team, 2016).

5.3.4 Model application

Two practical applications of the model were performed.

Yield forecasting

The model was applied to forecast yield anomalies during the growing season up to two months before harvest. We clipped the last one or two months, respectively, from the MIRCA2000-defined growing season and calculated all weather variables based on this reduced season. Afterwards the model was trained on the reduced weather data set, relating yield anomalies to weather anomalies observed up to one or two months before harvest. The one-out-of-sample performance of this reduced model is then a measure for its forecasting capacity.

Yield effects from temperature warming

Effects of moderate warming were calculated as a model application case. Temperature in every *second* growing season of the AgMERRA climate was raised by 0.9 or 1.4 °C, corresponding to the difference between the 0.6 °C of warming already present in 1986-2005 (Schleussner et al., 2016) and current climate change targets of 1.5 or 2 °C. Differences in warming over land and ocean (IPCC, 2013) were neglected. Precipitation and radiation were not modified since we assume stochastic changes with mean zero for this temperature range (IPCC, 2013). Differences in CO_2 concentrations would be relevant for absolute yields, but were not considered due to rather minor changes (plus ~30 or 60 ppm for 0.9 or 1.4 °C warming, respectively, compared to 1980-2010 average concentrations; IPCC (2013)). The CO_2 increase of ~60 ppm during the historical period is not relevant for this application when assuming a similar increase in the warmed period – first differences cancel the trend in both time series. Yield anomalies were predicted with coefficients estimated from unmodified climate and exogenous variables from the artificial climate data. Grid-cell yield time series were nationally aggregated without weighting. The first-difference approach allows interpreting yield changes between adjacent years as effects of temperature increases. Yield changes (unmodified to modified and modified to unmodified years, with inverted signs) were averaged and the logarithm removed. A temperature change of 0 °C was used for deriving normalization constants with which all other yield changes were multiplied. Uncertainty of predictions u was calculated by adding RMSE of the one-out-of-sample model ($RMSE_{O1}$) and variance of the temperature-modified yield time series (eq. 5):

$$u = \sqrt{(RMSE_{O1})^2 + Var(mod.time\ series)} \tag{5}$$

5.4 Results

5.4.1 Results for the contiguous US

The model had a substantial capacity for explaining and predicting yield anomalies. Yield anomaly time courses for USDA-based models are shown in Fig. 2. Results for each of the eight crop-yield data set combinations are displayed in Tab. 1. All grid cells where the specific crop is grown are included. Either unweighted or weighted aggregation was used, decided on the higher R^2_{O1} for each crop individually. Time series for US regions are provided in SI Fig. S11. A performance comparison of different model specifications is provided in SI Fig. S6. All statistical tests indicated that the OLS model estimation is adequate (SI section 4).

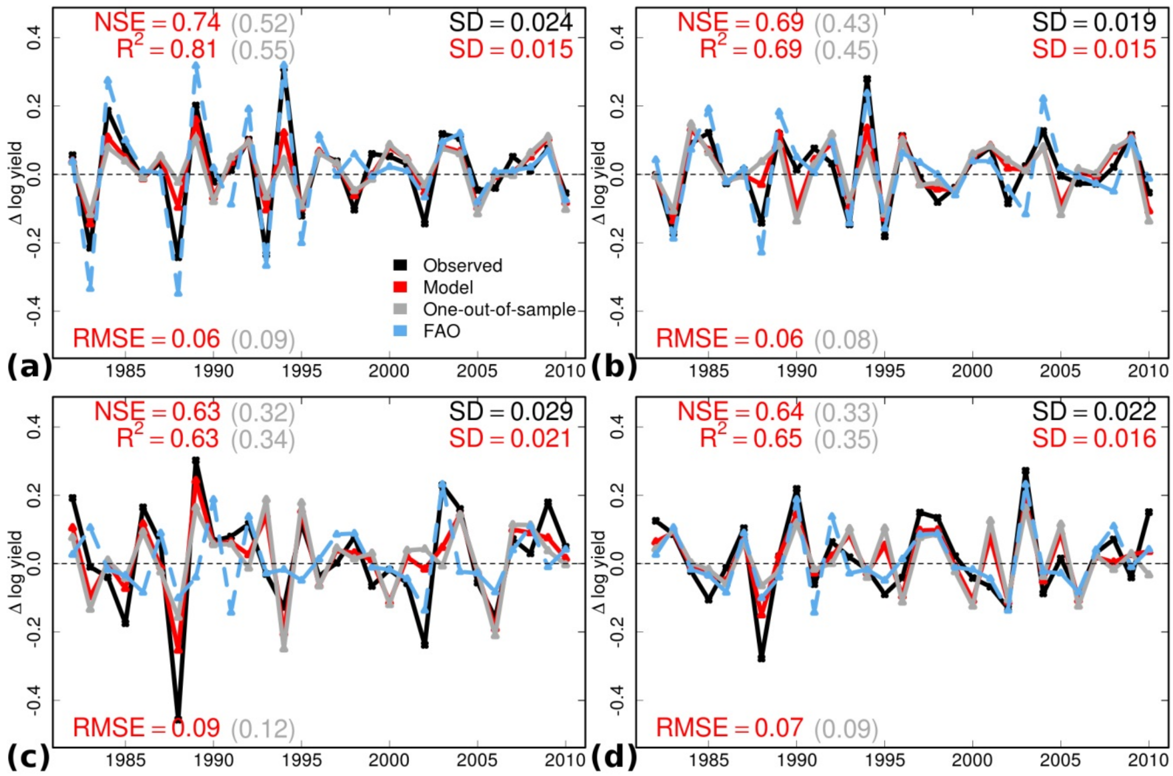


Fig. 2: Observed and modeled time series of national US yield anomalies for maize (a), soybeans (b), spring wheat (c) and winter wheat (d). Black lines are anomalies of reported USDA yields, red lines are anomalies predicted by the model trained on the full data panel, gray lines are anomalies predicted from one-out-of-sample models, and blue dashed lines are FAO yield anomalies. Data points were 56,092, 38,373, 21,291 and 58,877 for maize, soybeans, spring and winter wheat, respectively. Numbers in plots are performance measures and standard deviation (SD); colors of numbers correspond to the respective anomaly series. Modelled and FAO yield anomalies were significantly ($p < 0.05$) correlated for maize (Pearson's r = 0.87), soybeans (0.69) and winter wheat (0.68), but not for spring wheat (0.13), since FAO yields combine spring and winter wheat.

The model achieved at least two thirds of explained variance (R^2) and a robust (i.e. at least 25%) one-out-of-sample performance (R^2_{O1}) for all four crops with USDA data. Extremely low yields, like those occurring during the US heat and drought wave in 1988 for maize and wheat, were captured by the model, though not in full magnitude. For the two wheat types, yield loss quantities over the whole time series were comparable between model and observations, and for winter wheat also between one-out-of-sample model and observations. The set of three years of most negative yield anomalies (bottom decile) was equal for observed and modeled time series in 7 out of 12 cases. The observed top decile was captured in 8 out of 12 cases. For the one-out-of-sample predicted yields the correspondence for the bottom decile was less accurate with only 3 out of 12 cases. The direction of change and the sign of modeled anomalies matched with the input data for all crops, with only few exceptions.

Tab. 1: Model performance for eight crop-yield data set combinations in the US. Columns are crop, yield data set, application of land-use weighted aggregation, Nash-Sutcliffe efficiency (NSE), explained variance of the modeled (R^2) and one-out-of-sample time series (R^2_{O1}), out-of-temperature and out-of-precipitation correlation (R^2_{OOT} and R^2_{OOP}) and the share of grid cells for which the model is significant ($p < 0.05$).

Crop	Yield data	Weighted Aggregation	NSE	R^2	R^2_{O1}	R^2_{OOT}	R^2_{OOP}	Significant Cells
Maize	USDA	No	0.74	0.81	0.55	0.31	0.11	51 %
	GGYD	No	0.70	0.92	0.59	0.08	r<0	47 %
Soybeans	USDA	No	0.69	0.69	0.45	0.38	0.02	60 %
	GGYD	Yes	0.60	0.72	0.18	r<0	r<0	24 %
Spring wheat	USDA	No	0.63	0.63	0.34	0.28	0.42	52 %
	GGYD	No	0.61	0.73	0.32	r<0	0.34	48 %
Winter wheat	USDA	Yes	0.64	0.65	0.35	0.33	0.28	50 %
	GGYD	Yes	0.55	0.91	0.26	0.00	0.00	10 %

The model performed differently for different crops, judged by R^2_{O1}. The regression method, variable set or difference method influenced model performance (SI Fig. S6). Unweighted aggregation was better for maize, soybeans (except GGYD soybeans where R^2_{O1} was low) and spring wheat, but disfavored for winter wheat. Model performance differed between the two yield data sets. Although R^2 values were similar or higher for GGYD yields, R^2_{O1} values with GGYD data (Tab. 1, SI Fig. S6) were lower in three of four cases. Differences between R^2 and R^2_{O1} were thus higher for GGYD yields. STSM models showed, on average over all crops and specifications, slightly higher R^2 and R^2_{O1} values than PDM models (SI Fig. S6). R^2 and R^2_{O1} were correlated for USDA yields ($r = 0.97$, $p = 0$, $n = 24$), but not GGYD yields ($r = 0.29$, $p = 0.17$, $n = 24$). NSE and R^2 showed larger differences for GGYD than USDA yields. Thus the model's explanatory power was not an indicator for the model's projective power with GGYD yields. The out-of-temperature and out-of-precipitation performance (where six anomalies were omitted for training) was lower than the one-out-of-sample performance. All out-of-temperature values with USDA yields are, nevertheless, above 0.25, thus higher than expectable by chance (corresponding to $r = 0.5$). One-out-of-sample performance in the three warmest years is hardly different from modeled values. Out-of-precipitation values are above 0.25 only for wheat.

The explained variance varied spatially (Fig. 3). There was a substantial fraction of grid cells where the model was able to capture yield variability to a large (green shades) or an intermediate extent (yellow shades). However, there were also several regions where the model failed to capture variability (red shades). For all crops these were located in areas where yield variability was lower compared to other regions. In regions with substantial yield variation (coefficient of variation CV, defined as standard deviation over mean, is larger than 15%) the model achieved a higher R^2 more often (SI Fig. S10; SI Tab. S2). There was a moderate fraction of grid cells (11-27%) that exhibited low yield variability and was not well explained by the model.

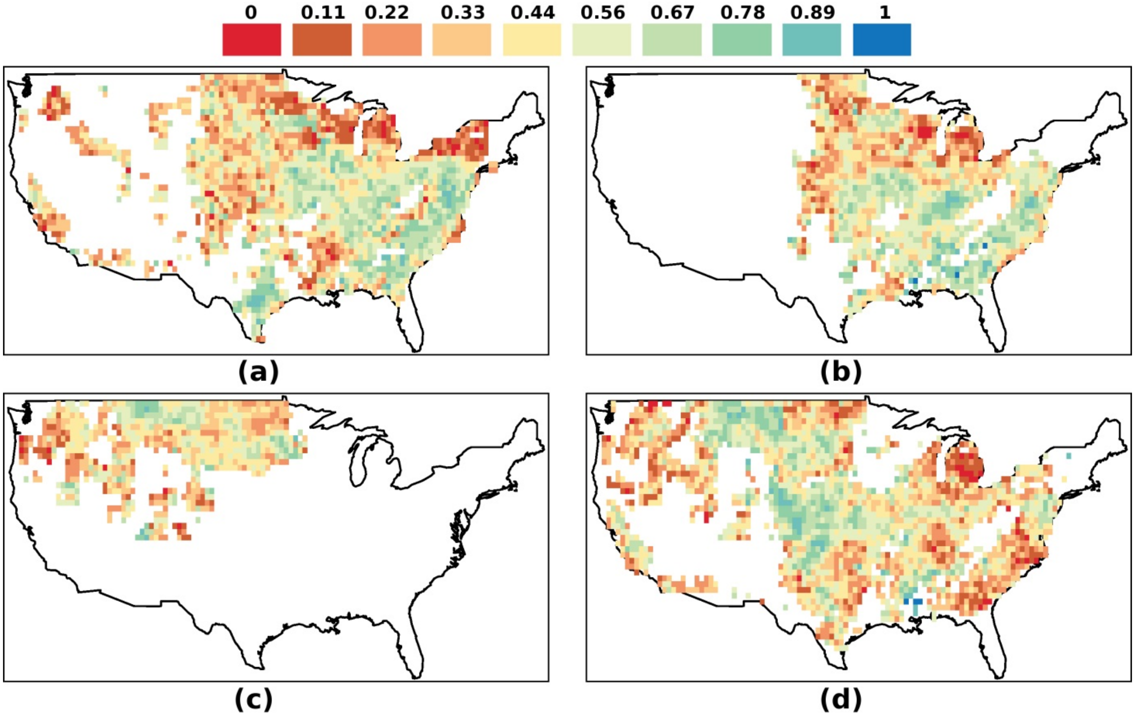


Fig. 3: Explained variance of yield anomalies due to weather anomalies (R^2, color map on top) for maize (a), soybeans (b), spring wheat (c) and winter wheat (d) with USDA yields. White regions have no cropping area.

Model coefficients indicated crop-specific patterns of weather influence. The influence of coefficients depended on the crop, but was independent from the estimation method (Fig. 4). All STSM coefficient means except two were significantly different from 0 (t-test at 95% confidence level). For all crops a high PET in the reproductive period was clearly negative. Precipitation was positive for summer crops during the vegetative period and for soybeans and winter wheat also during the reproductive period. For spring wheat and maize too much precipitation during the reproductive period was negative. Normalized solar radiation was negative for maize and soybeans (vegetative period), but strongly positive for spring and winter wheat. Any day above 32°C was damaging for all crops (not significant for winter wheat), whereby maize was most affected. Days below -15°C or 0°C, respectively, were damaging for all crops, but did not occur during the spring wheat growing season. There was a marked difference of coefficient values between the two yield data sets (USDA, GGYD). This was the case for STSMs (SI Fig. S7) and PDMs (SI Fig. S8).

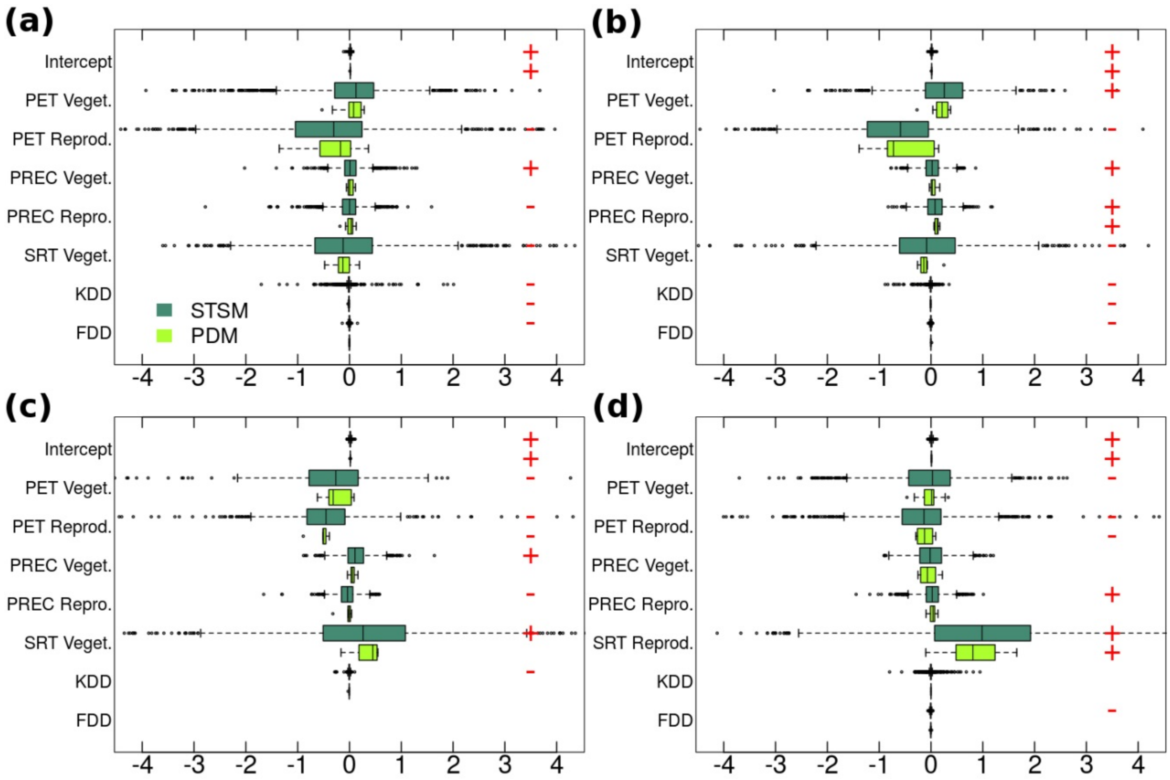


Fig. 4: Coefficient comparison for STSM and PDM model estimation for maize (a), soybeans (b), spring wheat (c) and winter wheat (d) with USDA yields. Blue boxes show coefficients with STSM estimation (estimated for each grid cell), while green boxes show PDM coefficients (estimated for each climate region). The band inside each box is the median, while boxes represent 25% and 75% quantiles. Whiskers are defined as the maximum and minimum as long as both values are within the 1.5 interquartile range from the median. Otherwise the last points in this range are shown with whiskers and outliers are depicted as points. Red +/- symbols indicate a mean significantly larger/lower than 0 (t-test at 95% confidence level).

Coefficients varied between climate regions (

Fig. 5**)**. A high PET during the vegetative season was positive for maize yield in the northern climate zones, but negative in the south. Vegetative PET was positive everywhere for soybeans. For spring wheat a high PET was negative everywhere except the northwest. For winter wheat a high PET during the reproductive season was positive only in the northeast, but negative elsewhere. The effect of precipitation did not show pronounced regional diversity: it was positive in most regions for all crops, with few exceptions. Elevated SRT during the vegetative period had a positive effect on maize yields in mid and western states, but not elsewhere. Enhanced SRT was negative for soybeans in all regions. For spring wheat, by contrast, higher SRT was positive everywhere except the northwest. For winter wheat more SRT had positive effects during the reproductive period in almost the whole US, with a positive gradient to the southeast. Days above 32°C were harmful everywhere for maize, spring and winter wheat (-2 to -4% yield loss for each day).

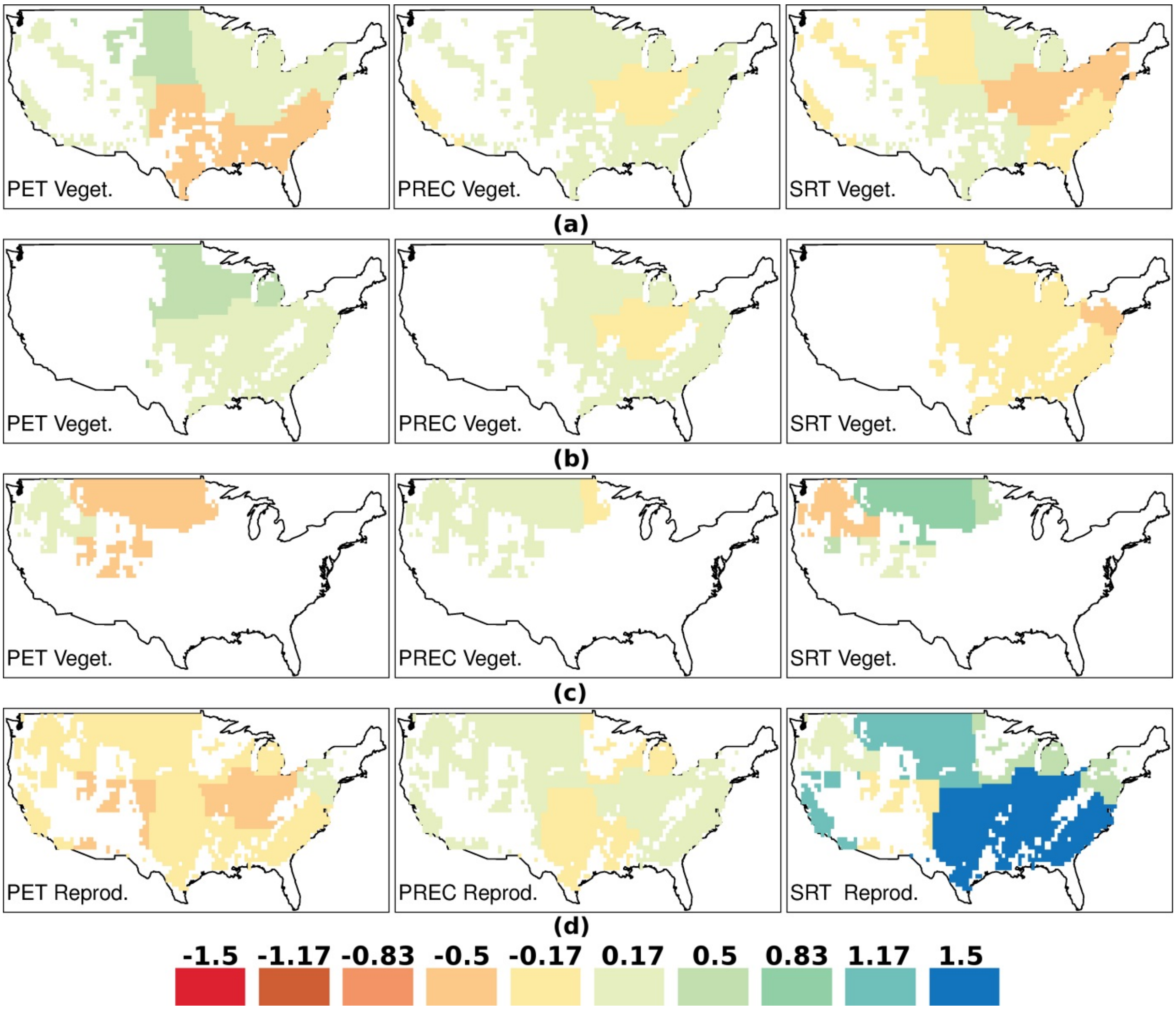


Fig. 5: Estimated coefficients for USDA yields. Rows are maize (a), soybeans (b), spring wheat (c) and winter wheat (d). Coefficients were estimated with STSM regression and aggregated from grid cells to climate regions. From left to right the coefficients are PET in vegetative (maize, soybeans, spring wheat) or reproductive (winter wheat) season, precipitation and SRT in the same seasons, respectively. Color map is shown at bottom.

A mapping sensitivity test, where climate, land-use and growing seasons were interpolated from grid cells to counties rather than yields from counties to grid cells, showed similar or slightly higher R^2 (0.82, 0.74, 0.65 and 0.68 for maize, soybeans, spring and winter wheat, respectively) and R^2_{O1} values (0.61, 0.55, 0.34 and 0.30). We kept the mapping of yields to grid cells, though, to maintain a common framework for both yield data sets.

5.4.2 Results for global main producers

The model explains more than two thirds of yield variance in main producer countries. The robust out-of-sample performance in the US supported an extension of the evaluation to other main producers (SI Tab. S3; Fig. 1). Only GGYD yields could be used as generally available source here. Nationally aggregated GGYD yield anomalies mostly corresponded well with FAO yield anomalies (SI Fig. S12), motivating the usage of this data set. The performance (R^2 and R^2_{O1}) for all crops is displayed in Fig. 6. The explained variance among main producers, weighted by total production, was 84%, 72%, 71% and 71% for maize, soybeans, spring and winter wheat, respectively. The weighted average one-out-of-sample performance was 42%, 22%, 33% and 15%. The cumulative production share (within the main producers) of nations which achieved an R^2_{O1} of at least 25% is 64%, 18%, 68% and 30% for maize, soybeans, spring and winter wheat, respectively. Analyses with PDM estimation led to similar, though slightly lower performances (SI Fig. S14). Calculating aggregated model performance as average performance over all grid cells in a country, rather than by correlating previously aggregated yield time series, resulted in lower model performances: mean R^2 [R^2_{O1}] STSM values over countries were 0.47 [0.18], 0.44 [0.15], 0.48 [0.19] and 0.36 [0.10] for maize, soybeans, spring and winter wheat. This aggregation effect, as discussed in Gornott and Wechsung (2016) for Germany, was thus confirmed globally.

Yield time series for selected main producers can be found in the supplement (SI Fig. S13). Mean performance was best for maize (highest R^2 and R^2_{O1}). While R^2 was similarly high for soybeans, the R^2_{O1} was rather low (22%). For winter and spring wheat the model achieved equal mean R^2, while mean R^2_{O1} was substantially higher for spring wheat. There was no obvious influence of harvested area, length of yield time series, share of rainfed agriculture, mean yield level or standard deviation on model performance. Countries where GGYD yields were constructed from subnational data (Table S1 in Iizumi et al. (2013b)) tended to have a larger R^2_{O1}, but not significantly. There are some notable discrepancies between R^2 and R^2_{O1}, especially for winter wheat: for example in India or Egypt an R^2 of 0.93 and 0.73, respectively, was accompanied by an R^2_{O1} of 0.04 and 0.03. In both cases, this discrepancy is due to extreme yield values captured by the model, but not the one-out-of-sample model (data not shown). If these extremes are removed, R^2_{O1} increases to 0.16 and 0.22, respectively. Differences between R^2 and R^2_{O1} are generally due to an out-of-sample time series which is less variable and captures fewer extreme values than the modeled time series.

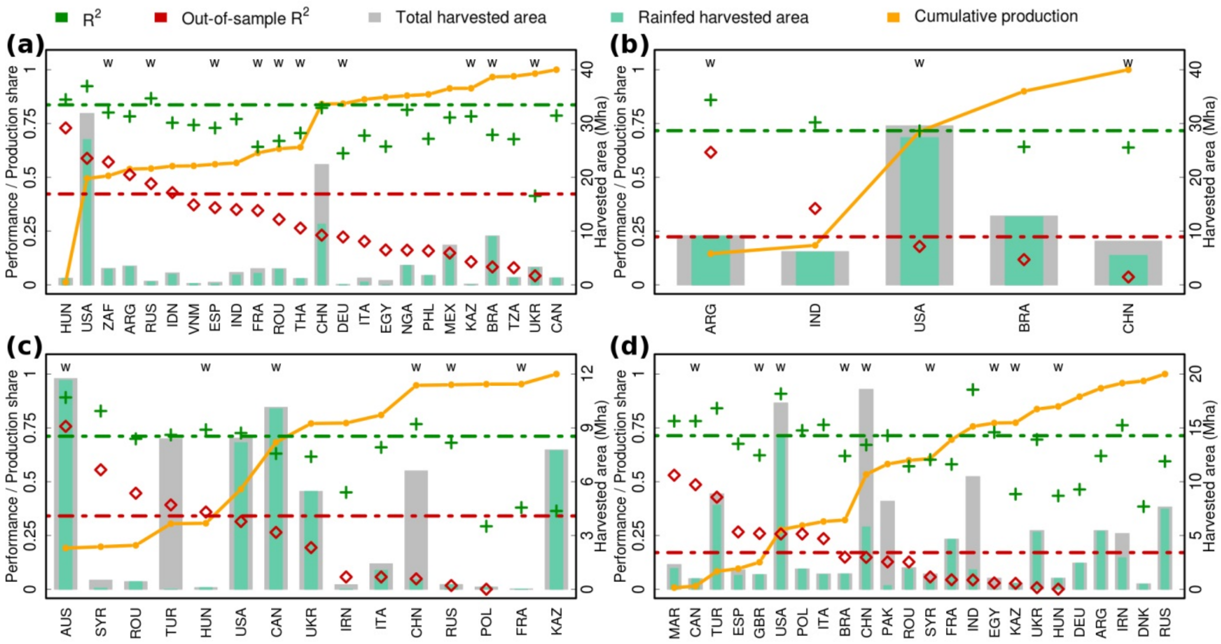


Fig. 6: Performance of STSM models in main producing countries for maize (panel a), soybeans (b), spring wheat (c) and winter wheat (d). Countries are ordered by descending R^2_{O1}; three-letter codes are provided in SI Tab. S3. Green crosses mark R^2 and red diamonds R^2_{O1} values (left y axis). The mean R^2 and R^2_{O1} over all main producers, weighted by production, are indicated with dashed green and red lines, respectively. A "w" above countries indicates that the displayed R^2_{O1} value is achieved when including land-use weighting. Gray and blue bars denote total and rainfed harvested area in Mha, respectively (right y axis). The orange line denotes cumulative production share among main producers (left y axis).

Yield data quality influences the detection of weather influences. There was a marked difference in model performance when using either reported sub-national yield data or gridded yield data derived from remote sensing. R^2_{O1} values for USDA data were 55%, 45%, 34% and 35% for maize, soybeans, spring and winter wheat, respectively, while for GGYD data these were 59%, 18%, 32% and 26%, thus lower except for maize (Tab. 1). This difference was also visible for Germany, Russia, Burkina Faso, Tanzania and Brazil (SI Tab. S4).

The average explained variance over all main producing countries and crops was 41.8% with GGYD yields. This was slightly higher than the 32-39% which have been found by Ray et al. (2015) with reported data. For maize the average R^2 was 44% with our model, compared to 39% in Ray *et al.*, and for soybeans it was 42%, compared to approx. 35%. For wheat (average over spring and winter) it was 42% with our model, compared to 35%.

Yield anomalies are forecasted with high accuracy within the growing season in several countries. The model was used for a simple forecasting of yields up to two months before harvest. The results for countries with reported yields are shown in Fig. 7, for all main producers using GGYD yields in SI Fig. S15. In all but five (out of 14) cases the one-out-of-sample performance is equal or even higher than the standard model when omitting the last month of the reproductive season for training and prediction. In seven cases this holds also when omitting the last two months. In ten cases yield anomalies can be predicted better than by chance ($R^2_{O1} > 0.25$) two months before harvest, and in six cases this prediction accuracy is more than 50%. When using GGYD yield data, 25 of 63 cases can be predicted with at least 25% accuracy two months before harvest (representing 4-86% of global production depending on the crop), and in six cases with 50% accuracy (representing 0-51% of global production).

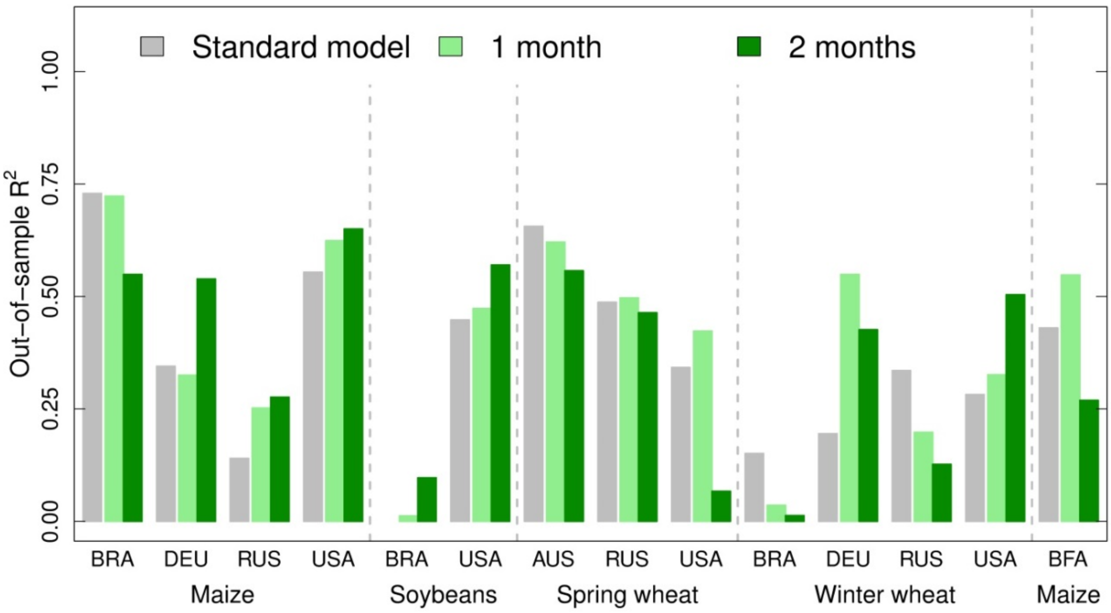


Fig. 7: Capacity of the model for yield forecasting within the growing season, using only reported yield data. The one-out-of-sample performance R^2_{O1} is shown. Gray bars are the standard model with full growing season used for training and prediction. Green and black bars show performance when withholding one or two months, respectively, for training the model and predicting yield anomalies out of sample. Burkina Faso (BFA) is not a main producer and therefore plotted off set.

Mean warming suggests negative yield effects. When increasing temperatures by 0.9 or 1.4 °C above the 1980-2010 average, yields are predicted to lose 3-18% (excluding Australian wheat and Brazilian soybeans) in comparison to reported yield data (Tab. 2). Results for Russia had high uncertainties due to large $RMSE_{O1}$ values and standard deviations. Projections based on GGYD yields were not performed due to low R^2_{OOT} scores (Tab. 1).

Tab. 2: Yield effects (as fraction of average historic yields) of artificial temperature increases, using only reported yield data. Fractions were normalized with T+0 offset. Values in brackets are uncertainty measures u (+/-) of the fraction according to equation 5.

Crop	Country	T +0.9 °C	T +1.4 °C
Maize	USA	0.96 (0.07)	0.95 (0.07)
	Russia	0.88 (0.87)	0.85 (0.86)
	Brazil	0.97 (0.19)	0.95 (0.20)
	Germany	0.96 (0.09)	0.94 (0.09)
	Burkina Faso	0.95 (1.00)	0.94 (1.00)
Soybeans	USA	0.97 (0.16)	0.96 (0.17)
	Brazil	1.00 (0.12)	1.00 (0.12)
Spring wheat	USA	0.95 (0.16)	0.92 (0.17)
	Australia	1.05 (0.71)	1.07 (0.74)
	Russia	0.89 (0.77)	0.84 (0.83)
Winter wheat	USA	0.97 (0.07)	0.95 (0.07)
	Russia	0.88 (0.72)	0.82 (0.78)
	Germany	0.95 (0.06)	0.92 (0.07)
	Brazil	0.89 (0.32)	0.85 (0.36)

5.5 Discussion

We have applied a semi-empirical regression model to estimate weather influences on yields of maize, soybeans, spring and winter wheat. The model achieves good performance in explaining and predicting inter-annual yield variation in the US. For all main producer countries a high average explanatory power but varying out-of-sample prediction capacity is attained. The model shows medium to high accuracy for yield anomaly forecasts during the growing season up to two months before harvest. An application of the model with artificially increased temperatures suggests negative effects of moderate warming on crop yields.

5.5.1 Modeling yield anomalies in the US

The fraction of explained yield variation was at least two thirds and the one-out-of-sample yield prediction accuracy achieved 34-55%. The model also achieved a quantitative reproduction of negative yield anomalies in most cases, which is of particular importance when studying non-linear economic responses. When validating the model in the warmest or driest years its out-of-sample capacity is better than 25% in six of eight cases (Tab. 1, USDA).

Explanation (R^2) and projection (R^2_{O1}) capacity were strongly different (up to 0.65) in some cases, and more so for GGYD yields (SI Fig. S6), underlining that both model fit and out-of-sample performance should be considered when evaluating the quality of a model (Holzkämper et al., 2015, Landau et al., 2000, Refsgaard et al., 2013). Differences between NSE and R^2 values could be due to an over-proportional influence of outlier values or scale effects on the NSE.

The different out-of-sample performance of the model with USDA and GGYD yield data, in particular for soybeans and winter wheat, suggests several uncertainties of the gridded yield data. First, the combination of reported yields with remote sensing data and growing season modeling might not be apt for winter crops as these are more easily mixed with other vegetation. Second, the time series of the GGYD data is shorter by six years, leaving less data for out-of-sample estimations. Yet a regression with USDA yields in the shorter GGYD time frame produced similar results as with the full range (data not shown), thus the shorter time series alone is unlikely to explain different performances. Third, the equal or higher average R^2 with GGYD yield data (SI Fig. S6) could possibly result from an implicit consideration of weather influences in the GGYD data set or the fitting of the model to more extreme values which arose in the GGYD construction but are not necessarily caused by weather. A misestimation of the true weather influence with our model would ensue. FAO yields, which are used in GGYD construction to calibrate remote sensing data, are often combined from reported and estimated data, adding a further layer of uncertainty. Fourth, yield variability from small plot sizes, in particular in developing countries, could be flattened at the coarse aggregate scale and thus blur weather influences. Fifth, GGYD yields showed lower CVs than USDA yields (except spring wheat, SI Tab. S2). This may explain the larger differences between R^2 and R^2_{O1} for GGYD yields, as low CVs together with shorter time series can lead to high correlations, but instable models i.e. a low R^2_{O1}. Similar differences in model performance between observed and remote sensing-derived yields in other nations (SI Tab. S4) further support our conclusions.

The geographical variation of model performance could have several causes. Different management techniques eliminate different shares of weather influence on crop yield. In particular irrigation, which is more prominent in the Western US (Schlenker and Roberts, 2009), marginalizes the effect of precipitation and also temperature (Lobell and Bonfils, 2008, Schauberger et al., 2017). This is underlined by a lower model performance in this region (Fig. 3). Thus, a low explanatory power might reflect a limited influence of weather on yields, as our model only detects weather impacts. Other reasons could include unconsidered, indirect weather influences (e.g. pests or diseases), errors in observations or aggregation effects. This may also explain the substantial share of grid cells with high yield variability but low explanatory power (SI Tab. S2). Low yield variability is difficult for any model to capture. Combined analysis of yield variation and model explanatory power reveals that areas with low yield variability are more likely to have a lower R^2 (SI Tab. S2, SI Fig. S10). Areas with a high USDA yield CV, by contrast, have equal shares of high and low explained variance. Uncertainties introduced by interpolating yield or weather statistics could destroy their associations (Hansen and Jones, 2000). A comparison of our results using GGYD data to the global study by Ray et al. (2015), using reported data, revealed a similar or larger share of grid cells with substantial yield variability but unsatisfactory explained variance ($R^2 < 0.45$) in Ray et al. Our results suggest, again, that yield variability in many agricultural areas is influenced by more factors than only weather. These could include changing land-use patterns (Olmstead and Rhode, 2011), economic influences like fertilizer usage or stressors like ozone or pests.

The estimated coefficients and their geographical distributions agree with expectations. Maize reacted negatively to a high PET in the reproductive season and to very hot days (KDD) in particular in warmer regions – which agrees with previous findings (Lobell et al., 2013, Schlenker and Roberts, 2009). This is contrary to expectations that C_4 crops would not experience much damage from mild heat (Sage and Kubien, 2007), but is likely due to water stress prior to direct heat damages (Schauberger et al., 2017). This effect also explains the higher model performance for maize and soybeans in the South, where water stress is more dominant. PET in the vegetative season and solar radiation affected maize positively only in cooler regions, confirming previous studies (Long et al., 2006, Rötter and Van de Geijn, 1999). Precipitation effects seem limited, though vegetative precipitation was usually positive. This conforms with a larger water demand of maize during the vegetative season (Hlavinka et al., 2009). The relatively low precipitation coefficient values, despite its prominent importance (Barnabas et al., 2008, Troy et al., 2015), are due to comparably high and strongly varying input values (Gornott and Wechsung, 2016, Lobell et al., 2013).

Differences in C_3 (soybeans, wheat) and C_4 (maize) photosynthesis efficiencies (Long et al., 2006, Rötter and Van de Geijn, 1999) are reflected in a lower positive effect of SRT for maize. KDDs were less negative for winter wheat than for maize, since these hardly occur during the growing season – winter wheat is usually harvested before heat waves build up. A higher PET in the reproductive cycle was more detrimental than a higher PET in the vegetative cycle of either winter wheat or maize due to a more developed canopy. This also applies to precipitation effect differences between the reproductive winter wheat and the vegetative maize cycle. The model performance was low for all crops in the Northwest and only slightly higher in the East North Central region. These regions seem more stable against weather fluctuations.

Six independent statistical tests indicated that our OLS estimation approach is applicable. Quadratic variables would not improve the model fit although this technique is often used to capture non-linear influences (Lobell et al., 2011, Ray et al., 2015). Autocorrelation occurring in many grid cells (SI Fig. S9) points to periodically occurring yield variability, which might lead to an underestimation of standard errors with OLS. But this autocorrelation is due to autocorrelation in the raw yield data (55%, 32%, 31% and 37% of grid cells for maize, soybeans, spring and winter wheat, respectively, at 95% confidence level with a Ljung-Box test) and the first difference approach which produces correlated yield differences. Therefore we assume it as unproblematic for our analysis. The nationally aggregated time series was weakly autocorrelated for soybeans and winter wheat and not autocorrelated for maize and spring wheat.

When calculating yield variability on spatially aggregated level, a land-use weighting is usually applied to capture spatially divergent contributions to agricultural production. But model performance was better with unweighted yields except for winter wheat, whose growing area is less concentrated (SI Fig. S3). Land-use patterns can be considered as an indirect function of climate since crops more favored by a certain climate also tend to have more area share. Thus there is an implicit inclusion of land-use patterns in the estimated coefficients, which makes the weighting negligible when inspecting aggregated yield variability. The differences are not substantial in all cases, which further suggests that land-use weighting can be omitted. This is beneficial for model generalization since weighting is another level of uncertainty (Cohn et al., 2016, Porwollik et al., 2016).

The model only used monthly aggregated weather data as input. This is an advantage over models requiring daily weather input since monthly aggregates are the preferred output from climate models (Taylor et al., 2012) and are also less sensitive to outliers. The yield-anomaly approach of our model additionally eliminates any time-dependent systematic bias. It is therefore particularly apt for usage with data from climate models, which often require a bias correction before impact assessments (Hempel et al., 2013).

5.5.2 Application to main producers

The generally good correlation between GGYD and FAO yield anomalies (SI Fig. S12) allows us to interpret aggregated production from GGYD yields and MIRCA2000 areas as representative for main producing countries. The average R^2_{O1} was at least one third for maize and spring wheat. For soybeans and winter wheat average R^2_{O1} was low, which is likely due to shortcomings of GGYD data with these crops (see above and below). This is supported by the increased performance of the model when using reported yield data (SI Tab. S4).

More than half of the global maize and spring wheat production anomalies could be well explained by our model (R^2_{O1} at least 25%). This enables the usage of our model in global economic assessments. We assume this share to rise with more reported yield data.

Countries with a high predictive capacity of the model (R^2_{O1} above or around 50%) all have water-dominated yield variability, i.e. the majority of cultivated area being rainfed and a rather high alternation between deficient and sufficient precipitation. This suggests that the model particularly captures water-limiting signals, though this may be questioned by the low R^2_{OOP} with GGYD yields (Tab. 1). Wheat grown in Morocco and Turkey was classified as winter wheat due to its relatively long growing season (7-11 months) over the local winter, but is different from “classical” winter wheat grown in cooler nations where the crop experiences a vegetative pause over the winter. This could bias results towards lower R^2 values. The performance of our semi-empirical model, when run with reported yield data, was equal or superior to several previously applied statistical approaches (Iizumi et al., 2013a, Lobell and Field, 2007, Ray et al., 2015, Urban et al., 2012).

We analyzed GGYD yields as an alternative to reported yields in areas where such data are currently not available. But the model-based nature of the data set could introduce a bias to our results. The robust performance of the semi-empirical model in the US, Germany, Russia, Burkina Faso, Tanzania and Brazil allows its usage for identifying cases where GGYD yields presumably suffer from a construction bias. We speculate that an existing weather influence on crops could be blurred by GGYD construction steps and is therefore less detectable with our (or any weather-driven) model. R^2 and R^2_{OI} values are then further apart, for example due to GGYD-processing induced yield extremes that are uncoupled from weather influences. The less convincing results for soybeans and winter wheat match with the evaluation by Iizumi et al. (2013b) suggesting that GGYD data likely requires improvement for both crops. A remaining concern is whether estimating a statistical model from a data set (GGYD) and then using the same model to evaluate these data may confound conclusions. But two additional analyses confirm our assumption that estimation problems occur more likely when GGYD yields are involved. First, the out-of-sample performance of models trained on reported yields is clearly superior to models trained on GGYD yields (SI Tab. S4). Second, a cross-comparison of model-predicted yields with reported FAO data, but where the model has been estimated with GGYD data (SI Fig. S14), shows that there are discrepancies for all crops. Differences between predicted yields and FAO are usually smaller when using reported yields for training the model (dashed blue lines in Fig. 2). Nevertheless we esteem the unique ability of GGYD yields to cover all regions of the globe where subnational yield data are otherwise difficult to obtain. Usage of latest satellite data with more sophisticated land-use separation methods may reduce counter-factual error sources and thus increase the reliability of satellite-derived yield statistics (Iizumi and Ramankutty, 2016).

5.5.3 Yield forecasting and warming experiment

The model concept allows for a simple extension towards forecasting of yields few months before harvest. This study presents a first example application in this direction. The forecasting is robust (R^2_{OI} > 50%) up to two months before harvest in several major producing countries, but requires improvement in others, in particular for soybeans and winter wheat. The performance is thus comparable to previous approaches (Bolton and Friedl, 2013, Johnson, 2014, Sakamoto et al., 2014), but has been done here without any particular adaptation to country-specific conditions or model formulation. In several cases the reduced growing season leads to higher R^2_{OI} values than the full season. This could stem from three reasons. First, crop climatic requirements can be different in grain filling and maturity phase (Barnabas et al., 2008), which are not distinguished in our reproductive season and could lead to meaningless coefficients in the default model. Second, the growing season dates in MIRCA2000 could be wrong, leading to an improvement when omitting a too long part. Third, the vegetative and reproductive season split could be misplaced. These reasons will have to be investigated in further studies.

Again, the importance of high-quality input yield data for model training is highlighted: only then reliable within-season forecasts are possible, as evidenced by the lower performance with GGYD yields.

The forecasting scheme could be modified in two directions. Both require near-term monthly weather forecasts published, for example, by the NOAA (NOAA Climate Forecast, 2017). First, the full growing season can be used for training. In the season where yields should be predicted before harvest the missing part of the weather information is supplied by a near-term forecast. Second, both approaches can be combined: a reduced growing season, e.g. withholding the last two months of the season, is used for training. Yield predictions are then calculated for three or more months before harvest by supplying the missing weather information up to two months before harvest with near-term weather forecasts.

Predicting yields with counter-factual temperature increases is another model application case. The approach neglects CO_2 trends, variation of cofactors like precipitation and comes with high uncertainties (out-of-temperature performances in Tab. 1 and the u measure according to equation 5 provide a first, maybe too high estimate), which might mask effects. This could change if real climate scenarios were used including drifts in temperature extremes and precipitation. But impacts seem plausible in direction and magnitude compared to previous studies (Challinor et al., 2014, Giannakopoulos et al., 2009, Schleussner et al., 2016). The low R^2_{OOT} performance for GGYD yields underlines the importance of high-quality yield data when projecting future yields. The average decline in wheat yields, when averaged over spring and winter wheat at 0.9°C warming (Tab. 2), is 6% – in agreement with the results by Liu et al. (2016). Thus the semi-empirical model described here can be considered a fourth method next to the three methods considered therein.

The model scheme presented in this study is an open concept that can be extended to incorporate further weather or economic factors. The prediction of yields within the growing season is highly sought after for timely adaptation measures in management, storage or marketing. Our model will be further developed in this direction. The differential performance between observed and remote-sensing based yield data calls for better and publicly available yield data from statistical offices in all countries. These can aid in planning adaptation or evaluating, for example, agricultural micro-insurance schemes.

5.6 References

Barlow KM, Christy BP, O'leary GJ, Riffkin PA, Nuttall JG (2015) Simulating the impact of extreme heat and frost events on wheat crop production: A review. Field Crops Research, 171, 109-119.

Barnabas B, Jager K, Feher A (2008) The effect of drought and heat stress on reproductive processes in cereals. Plant Cell Environ, 31, 11-38.

Belsley DA, Kuh E, Welsch RE (1980) *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity,* New York, Wiley.

Bolton DK, Friedl MA (2013) Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. Agricultural and Forest Meteorology, 173, 74-84.

Challinor AJ, Watson J, Lobell DB, Howden SM, Smith DR, Chhetri N (2014) A meta-analysis of crop yield under climate change and adaptation. Nature Climate Change, 4, 287-291.

Cohn AS, Vanwey LK, Spera SA, Mustard JF (2016) Cropping frequency and area response to climate variability can exceed yield response. Nature Climate Change, 6, 601-604.

Conradt T, Gornott C, Wechsung F (2016) Extending and improving regionalized winter wheat and silage maize yield regression models for Germany: Enhancing the predictive skill by panel definition through cluster analysis. Agricultural and Forest Meteorology, 216, 68-81.

Croissant Y, Millo G (2008) Panel data econometrics in R: the plm package. J. Stat. Softw., 27, 1-43.

Di Paola A, Valentini R, Santini M (2016) An overview of available crop growth and yield models for studies and assessments in agriculture. Journal of the Science of Food and Agriculture, 96, 709-714.

Fao (2016) FAOStat, http://faostat3.fao.org/home/E.

Giannakopoulos C, Le Sager P, Bindi M, Moriondo M, Kostopoulou E, Goodess CM (2009) Climatic changes and associated impacts in the Mediterranean resulting from a 2 °C global warming. Global and Planetary Change, 68, 209-224.

Glotter M, Elliott J, Mcinerney D, Best N, Foster I, Moyer EJ (2014) Evaluating the utility of dynamical downscaling in agricultural impacts projections. Proc Natl Acad Sci U S A, 111, 8776-8781.

Gornott C, Wechsung F (2016) Statistical regression models for assessing climate impacts on crop yields: A validation study for winter wheat and silage maize in Germany. Agricultural and Forest Meteorology, 217, 89-100.

Hansen JW, Jones JW (2000) Scaling-up crop models for climate variability applications. Agricultural Systems, 65, 43-72.

Hatfield JL, Boote KJ, Kimball BA *et al.* (2011) Climate Impacts on Agriculture: Implications for Crop Production. Agronomy Journal, 103, 351-370.

Haude W (1955) Zur Bestimmung der Verdunstung auf möglichst einfache Weise. Mitteilungen des Deutschen Wetterdienstes, 11.

Hempel S, Frieler K, Warszawski L, Schewe J, Piontek F (2013) A trend-preserving bias correction - the ISI-MIP approach. Earth System Dynamics, 4, 219-236.

Hlavinka P, Trnka M, Semerádová D, Dubrovský M, Žalud Z, Možný M (2009) Effect of drought on yield variability of key crops in Czech Republic. Agricultural and Forest Meteorology, 149, 431-442.

Holzkämper A, Calanca P, Honti M, Fuhrer J (2015) Projecting climate change impacts on grain maize based on three different crop model approaches. Agricultural and Forest Meteorology, 214-215, 219-230.

Iizumi T, Ramankutty N (2016) Changes in yield variability of major crops for 1981–2010 explained by climate change. Environmental Research Letters, 11, 034003.

Iizumi T, Sakuma H, Yokozawa M *et al.* (2013a) Prediction of seasonal climate-induced variations in global food production. Nature Climate Change, 3, 904-908.

Iizumi T, Uno F, Nishimori M (2012) Climate Downscaling as a Source of Uncertainty in Projecting Local Climate Change Impacts. Journal of the Meteorological Society of Japan. Ser. II, 90B, 83-90.

Iizumi T, Yokozawa M, Sakurai G *et al.* (2013b) Historical changes in global yields: major cereal and legume crops from 1982 to 2006. Global Ecology and Biogeography, 23, 346-357.

Ipcc (2013) Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. (eds Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM), Cambridge, United Kingdom and New York, NY, USA, .

Johnson DM (2014) An assessment of pre- and within-season remotely sensed variables for forecasting corn and soybean yields in the United States. Remote Sensing of Environment, 141, 116-128.

Jones JW, Antle JM, Basso B *et al.* (2016) Brief history of agricultural systems modeling. Agricultural Systems.

Kilsby CG, Jones PD, Burton A *et al.* (2007) A daily weather generator for use in climate change studies. Environmental Modelling & Software, 22, 1705-1719.

Landau S, Mitchell RaC, Barnett V, Colls JJ, Craigon J, Payne RW (2000) A parsimonious, multiple-regression model of wheat yield response to environment. Agricultural and Forest Meteorology, 101, 151-166.

Liu B, Asseng S, Ewert F *et al.* (2016) Similar estimates of temperature impacts on global wheat yield by three independent methods. Nature Climate Change.

Lobell DB (2013) Errors in climate datasets and their effects on statistical crop models. Agricultural and Forest Meteorology, 170, 58-66.
Lobell DB, Bänziger M, Magorokosho C, Vivek B (2011) Nonlinear heat effects on African maize as evidenced by historical yield trials. Nature Climate Change, 1, 42-45.
Lobell DB, Bonfils C (2008) The Effect of Irrigation on Regional Temperatures: A Spatial and Temporal Analysis of Trends in California, 1934–2002. Journal of Climate, 21, 2063-2071.
Lobell DB, Burke MB (2010) On the use of statistical models to predict crop yield responses to climate change. Agricultural and Forest Meteorology, 150, 1443-1452.
Lobell DB, Field CB (2007) Global scale climate–crop yield relationships and the impacts of recent warming. Environmental Research Letters, 2, 014002.
Lobell DB, Hammer GL, Mclean G, Messina C, Roberts MJ, Schlenker W (2013) The critical role of extreme heat for maize production in the United States. Nature Climate Change, 3, 497-501.
Long SP, Zhu X-G, Naidu SL, Ort DR (2006) Can improvement in photosynthesis increase crop yields? Plant, Cell and Environment, 29, 315-330.
Luo Q (2011) Temperature thresholds and crop production: a review. Climatic Change, 109, 583-598.
Maurer EP, Hidalgo HG, Das T, Dettinger MD, Cayan DR (2010) The utility of daily large-scale climate data in the assessment of climate change impacts on daily streamflow in California. Hydrology and Earth System Sciences, 14, 1125-1138.
Mueller ND, Gerber JS, Johnston M, Ray DK, Ramankutty N, Foley JA (2012) Closing yield gaps through nutrient and water management. Nature, 490, 254-257.
Noaa Climate Forecast (2017) http://www.cpc.ncep.noaa.gov/products/forecasts/.
Olmstead AL, Rhode PW (2011) Adapting North American wheat production to climatic challenges, 1839 – 2009. Proc Natl Acad Sci U S A, 108, 480-485.
Porter JR, Gawith M (1999) Temperatures and the growth and development of wheat a review. European Journal of Agronomy, 10, 23-36.
Portmann FT, Siebert S, Döll P (2010) MIRCA2000-Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. Global Biogeochemical Cycles, 24, GB1011.
Porwollik V, Müller C, Elliott J *et al.* (2016) Spatial and temporal uncertainty of crop yield aggregations. European Journal of Agronomy.
R Core Team (2016) R: A Language and Environment for Statistical Computing. (ed Computing RFFS), R Foundation for Statistical Computing.
Rahmstorf S (2007) A Semi-Empirical Approach to Projecting Future Sea-Level Rise. Science, 315, 368-370.
Ray DK, Gerber JS, Macdonald GK, West PC (2015) Climate variation explains a third of global crop yield variability. Nat Commun, 6, 5989.
Refsgaard JC, Madsen H, Andréassian V *et al.* (2013) A framework for testing the ability of models to project climate change and its impacts. Climatic Change, 122, 271-282.
Ritchie SW, Hanway JJ, Benson GO, Herman JC, Lupkes SJ (1993) How a Soybean Plant Develops, Ames, Iowa State University of Science and Technology.
Rötter RP, Carter TR, Olesen JE, Porter JR (2011) Crop–climate models need an overhaul. Nature Climate Change, 1, 175-177.
Rötter RP, Van De Geijn S (1999) Climate change effects on plant growth, crop yield and livestock. Climatic Change, 43, 651-681.
Ruane AC, Goldberg R, Chryssanthacopoulos J (2015) Climate forcing datasets for agricultural modeling: Merged products for gap-filling and historical climate series estimation. Agricultural and Forest Meteorology, 200, 233-248.
Sage RF, Kubien DS (2007) The temperature response of C_3 and C_4 photosynthesis. Plant Cell Environ, 30, 1086-1106.
Sakamoto T, Gitelson AA, Arkebauer TJ (2014) Near real-time prediction of U.S. corn yields based on time-series MODIS data. Remote Sensing of Environment, 147, 219-231.
Sanchez B, Rasmussen A, Porter JR (2014) Temperatures and the growth and development of maize and rice: a review. Glob Chang Biol, 20, 408-417.
Schauberger B, Archontoulis S, Arneth A *et al.* (2017) Consistent negative response of US crops to high temperatures in observations and crop models. Nat Commun, 8, 1-9.
Schlenker W, Roberts MJ (2009) Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. Proc Natl Acad Sci U S A, 106, 15594-15598.
Schleussner C-F, Lissner TK, Fischer EM *et al.* (2016) Differential climate impacts for policy-relevant limits to global warming: the case of 1.5 deg C and 2 deg C. Earth System Dynamics, 7, 327-351.
Taylor KE, Stouffer RJ, Meehl GA (2012) An Overview of CMIP5 and the Experiment Design. Bulletin of the American Meteorological Society, 93, 485-498.
Troy TJ, Kipgen C, Pal I (2015) The impact of climate extremes and irrigation on US crop yields. Environmental Research Letters, 10, 054013.

Urban D, Roberts MJ, Schlenker W, Lobell DB (2012) Projected temperature changes indicate significant increase in interannual variability of U.S. maize yields. Climatic Change*,* 112, 525-533.
USDA (2015) USDA Quickstats, http://quickstats.nass.usda.gov/.
Wechsung F, Lüttger AB, Hattermann F (2008) Projektionen zur klimabedingten Änderung der Erträge von einjährigen Sommer-und Winterkulturen des Ackerlandes am Beispiel von Silomais und Winterweizen. In: *PIK Report.*
Wooldridge JM (2013) *Introductory Econometrics. A Modern Approach*, South Western Cengage Learning.

6 Covering smallholder farmers' weather perils – a crop model based insurance approach for Tanzania

Christoph Gornott[1*], Fred Hattermann[1], and Frank Wechsung[1]

[1] Potsdam Institute for Climate Impact Research (PIK)

[*] Corresponding author

6.1 Abstract

Weather-related yield losses endanger food security and inhibit the establishment of a resilient farming system for more than 30 million people working in the agricultural sector of Tanzania. If these losses were quantified, this information could be used for determining crop insurance claims to indemnify smallholder farmers in overall Sub-Saharan Africa and stabilize their incomes. These insurance solutions are addressed by the IPCC AR5, G7 leaders and COP21 as important tool to enhance Sub-Saharan Africa's resilience to climate change. Here, we develop a combined application of a process-based and statistical crop model and demonstrate that this approach significantly improves the yield assessment accuracy by 74% at district level. Furthermore, it allows to separate weather-related yield losses (covered by the insurance) from the management-related losses. Using our approach, we calculate that only 27% of the actual maize yield losses in Tanzania are directly attributable to weather. Considering this and the model uncertainty, the insurance premiums could decrease by one third – 71 million US$ p.a. (23 US$ ha^{-1}) – for maize production in Tanzania. Among other implementation components, our loss-determination approach can contribute to successfully implement index insurances for smallholder farmers and incentives lenders offering credits to these farmers.

6.2 Introduction

In Sub-Saharan Africa (SSA), crop yields commonly have high variability on a very low average yield level. This hinders smallholder farmers investing in agronomic management to stabilize and increase their crop yields and keeps them in the loop of poverty and food insecurity. Often, these smallholder farmers organize their risk management by self and community-based insurance strategies and thus, they do not have the financial capacity to adjust their agronomic management when extreme weather conditions strike (Carter, 2012; McIntosh et al., 2013). An improved agronomic management could contribute to stabilize smallholder farmers' incomes and make their agricultural production less vulnerable to weather extremes. Besides these farm-individual perils, widespread weather perils (termed *systemic risk*) strongly harm the agricultural sector as it was the case during the El Niño drought of

2014–15 and 2015–16 in eastern and southern Africa. Without proper risk transfer instruments, systemic risks make smallholder farmers highly vulnerable to crop yield losses (Müller et al., 2011). Index-insurance schemes have high potential as an adaptation strategy towards climate change and systemic weather perils (IPCC, 2014; Surminski et al., 2016), because they can stabilize smallholder farmers' incomes, prevent indebtedness, and indemnify their livelihoods. However, widespread implementation of such insurance schemes is hindered by uncertain and unreliable assessments of crop yield losses, notably for cropping conditions in SSA.

A successful yield insurance scheme should cope with systemic risks and should have low costs for both loss-determination and distribution of claim payoffs (termed *transaction costs*). Weather-index insurance schemes are designed to cover actual yield losses independently from the loss origin and thus, may oversimplify the weather yield relationship (Herbold, 2014). While weather-index insurances (see supplemental information (SI) S.3.2 for further information) have low transaction costs, but high basis risk due to imprecise coverage of actual losses (Leblois and Quirion, 2013), indemnity-based insurances for individual farms have accurate yield loss-determination, but often high transaction costs. Here, we combine the advantages of both insurances. Our modeling approach for an area-based yield insurance scheme (MAYIS) facilitates the identification of only weather-related yield losses, while non-weather-related yield losses are omitted. The non-weather-related perils are the yield impacts of socio-economic behavior or agronomic decisions, which could have been influenced by the insured farmer. If the insurer also bears the non-weather-related yield losses, this would cause a riskier behavior by farmers (moral hazard) and increase signed insurance policies by more vulnerable farmers (adverse selection). This might increase claims and hence, premiums and reduce the acceptance of the insurance (Conradt et al., 2015; Meze-Hausken et al., 2009; Shen and Odening, 2013).

In Tanzania, maize (*Zea mays* L.) is the most widely cultivated crop. Cropping conditions are characterized by high spatial and temporal heterogeneity (Ramirez-Villegas and Challinor, 2012; Rowhani et al., 2011). The average annual precipitation ranges in the south–west lowlands from 700 to 2,000 mm and in the northern semi-arid highlands from 400 to 700 mm. The monthly average temperature is between 18 and 28 °C throughout the year. Despite this favorable climate, the mean Tanzanian maize yield is rather low at 1.3 t ha^{-1}. Typically for SSA, yields are more often influenced by agronomic management than by weather impacts (Affholder et al., 2013; Lesk et al., 2016). In comparison, the impact of agronomic management on yield variability is smaller in regions with a high-input agronomic management (Ray et al., 2015). In SSA, a low and unbalanced fertilizer supply characterizes the agronomic management and represents the major yield limitations (Schlenker and Lobell, 2010; van der Velde et al., 2014). Besides weather and fertilization, several other factors influence maize yields (Moore et al., 2012). Among these factors are, notably, limited access to arable land (Iizumi and Ramankutty, 2015), labor, credits, markets, and technology (Herbold, 2014; Schlenker and Lobell, 2010), pests, weeds, and diseases (Rosenzweig et al., 2001), or fertilizer subsidies (Benson et al., 2012; Jayne et al., 2013; Sánchez, 2010).

Crop models can contribute to gaining insights about the impacts of weather, soil, agronomy, and socio-economy on crop yields. These insights of the crop models make it possible to separate the weather-related yield losses from the total yield losses. In most global and regional crop yield assessments, process-based (Asseng et al., 2013; Bassu et al., 2014; Folberth et al., 2012) and statistical models (Blanc, 2012; Ray et al., 2015; Rowhani et al., 2011; Schauberger et al., 2017) are used alternatively. Estes et al. (2013) show the advantages and weaknesses of these two model types for South-African crop yield assessments. Lobell et al. (2005) and Lobell and Burke (2010) separately use a statistical model to corroborate process-based yield assessments. Liu at al. (2016) and Lobell and Asseng (2017) show in an inter-comparison the similarities of both model types. However, to our knowledge no combined application of both model types has so far been published. Here, we use a process-based model to identify purely weather-attributable yield variability, while our statistical model captures the remaining non-weather-related yield variability. The ability of statistical models to account for non-weather-related impacts allows us to identify those yield impacts beyond the weather-attributable yield impacts. We combine the advantages of both model types to enhance the robustness of yield assessments and to integrate scarce observed yield data efficiently. This makes our approach suitable for other regions of SSA with also limited observed yield information.

6.3 Results and discussion

The combination of a process-based and statistical model increases the assessment accuracy of yield variability. The combined application of both model types significantly ($p < 0.01$, Fisher z-transformation, 796 observations) increases the reproduction of annual yield variability (Fig. 1). While the solely process-based assessment attains $r = 0.05^{NS}$ ($R^2 = 0.00$), the goodness of fit increases to $r = 0.86^{***}$ ($R^2 = 0.74$) for the combined assessment (Pearson correlation; $^{NS}\,p > 0.1$, $^{*}\,p \leq 0.1$, $^{**}\,p \leq 0.05$, $^{***}\,p \leq 0.01$) [all correlation coefficients and the corresponding R^2 are in SI Tab. S.1, S.2, and Fig. S.7]. Moreover, the out-of-sample validation achieves a correlation of $r = 0.38^{***}$ ($R^2 = 0.15$) and the corresponding statistical tests show that the model provides robust and valid results (see SI S.2.3.2 for details). However, the solely application of a statistical model to identify the weather-attributable yield variability significantly ($p < 0.01$, Fisher z-transformation) reduces the goodness of fit to $r = 0.77$ ($R^2 = 0.59$) for the estimation and to $r = 0.10$ ($R^2 = 0.01$) for the validation (see SI S.2.3.3 and Fig. S.8 for further details). This demonstrates that the information of the process-based and the statistical model are complementary. Since weather has often nonlinear and more complex impacts on crop yields, linear and log-linear statistical models are only limitedly able to capture the weather-attributable yield variability and therefore, often underestimate the weather impacts.

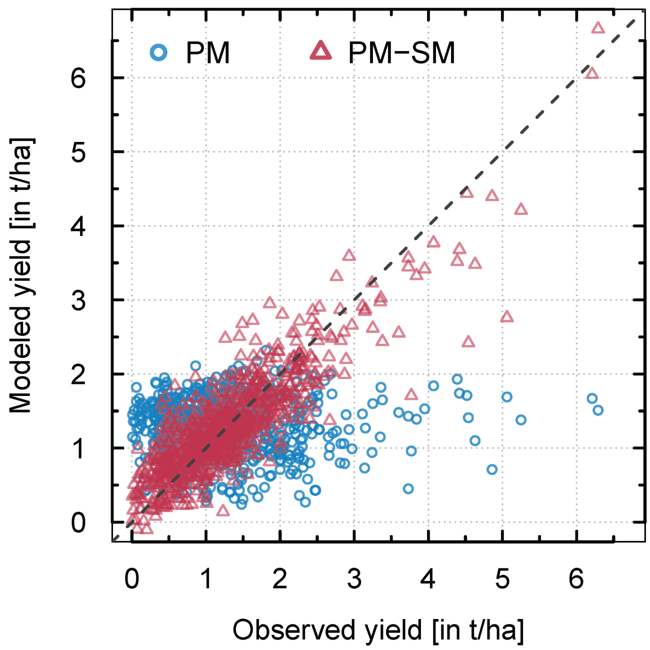


Fig. 1. Increase in goodness of fit due to the combined application of a process based (PM) and a statistical model (SM). The blue points show the accuracy of a solely application of the PM. The red points show the accuracy of a consecutive application of PM and SM (PM-SM).

Our process-based model satisfactorily captures the average national maize yield (modeled: 1.29 t ha^{-1} and observed: 1.27 t ha^{-1}) between 2003 and 2010. The process-based modeled yields show regional yield patterns of low and high yields similar to the observed district yields (Fig. 2). Aggregated to agro-ecological zones (see SI S.1.2 Fig. S.2), the modeled yields correlate spatially with the observed yields at $r = 0.57^{NS}$ ($R^2 = 0.32$). The semi-arid regions in the center and the north–eastern regions as well as the sub-humid regions in the south are clearly distinguishable in the modeled and observed yield maps. However, the annual yield variability is insufficiently reproduced by the process-based model for entire Tanzania, a result also found for other regions and process-based crop models (Müller et al., 2016). Nevertheless, the water scarce regions are reproduced with higher accuracy than the regions with sufficient water supply (SI Fig. S.5 and S.6). Since process-based models consider only a limited number of processes in their model set-up, they may neglect possibly relevant ones (Rötter et al., 2011). In particular, socio-economic impacts on agronomic management practices, which are important in SSA (Iizumi and Ramankutty, 2015; Ward et al., 2014), are usually not considered by process-based models. Our consecutively applied statistical model resolves the residual yield variability by using the non-weather variables *maize acreage*, *paid subsidies on crop production*, and *urea application* (see SI S.1.4 for further information). As a result, our combined modeling approach is able to reproduce the actual yield variability; this justifies the separation of weather and non-weather-related yield losses for the utilization in MAYIS.

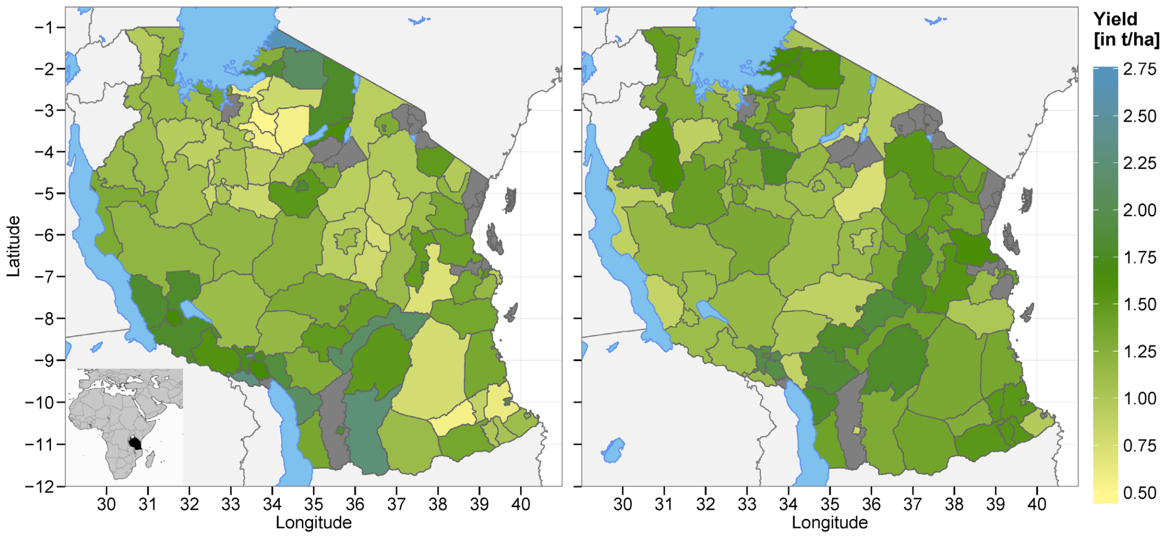


Fig. 2. Observed (left) and process-based modeled (right) average maize yields for Tanzanian districts in the period 2003-2010. No data is marked in dark gray.

Generally, our combined approach explains 74% of the total observed yield variability ($R^2 = 0.74$). Considering process-based models alone might cause them to be rejected as of limited use, if they fail to satisfactorily explain total yield variability. However, as shown, statistical models can be used to explain the remaining yield variability due to agronomic management and socio-economic factors. This demonstrates the general relevance and usability of process-based models. However, the aggregation of farm yields to the district level by the Tanzanian statistical office might have a filtering effect (Woodard and Garcia, 2008). If a more comprehensive yield dataset were available, it would allow to resolve this effect and to dissect the basis risk of index-insurances within districts. Nevertheless, our combined approach contributes a pragmatic solution to cover both agronomic management and socio-economic yield impacts in addition to weather impacts. This information is crucial for the acceptance of the area-based yield insurance scheme and justifies the usability of the weather-attributable yield variability for insurance purposes. Moreover, the process-based model in our combined approach allows assessments of yield losses for changed agronomic management practices and altered weather conditions, which are in agreement with plant-physiological processes.

Indirect weather-triggered effects are negligible for Tanzania. Besides both impact factor groups, we also investigate whether indirect weather-triggered effects (like pests and diseases) explain the remaining yield variability. Similar to the combination of the two model types, we estimate a consecutive weather-driven statistical model with the residuals of the non-weather-driven statistical model as the endogenous variable (see SI Fig. S.4). The weather-driven statistical model explains the residual yield variability by *precipitation*, *vapor pressure deficit*, and *solar radiation* of the district-specific growing season. This consecutive weather-driven statistical model explains yield variability with $r = 0.92^{***}$ ($R^2 = 0.84$). However, the validation decreases from $r = 0.38^{***}$ ($R^2 = 0.15$) [only process-based and non-weather-driven statistical model] to $r = 0.33^{***}$ ($R^2 = 0.11$) [process-based, non-weather, and indirect weather-triggered statistical model]. In a further step, we remove the non-weather-driven statistical model. Considering only the indirect weather-triggered effects significantly ($p < 0.01$, Fisher z-transformation) reduces goodness of fit to $r = 0.78^{***}$ ($R^2 = 0.60$) for the estimation and to $r = 0.04^{NS}$ ($R^2 = 0.00$) for the validation, respectively. Hence, we conclude that the indirect weather-triggered effects do not contribute model robustness. In the following we only consider the non-weather-related impacts to explain the residual yield variability. The results indicate that indirect weather-triggered impacts only have a minor influence on crop yields at the district scale. However, since process-based crop models are calibrated to field trials with a prevalent pest, disease and weed pressure, it is possible that these impacts are already implicitly included in our model. This could be the reason for the indirect weather-triggered impacts appearing insignificant. If there were a significant and robust influence of these indirect weather-triggered effects, it would be allocated to the weather-related part (because of the correlation with the weather) and thus, be indemnified.

Weather-related yield losses constitute only one-third of total maize yield losses in Tanzania. Crop insurance help stabilize smallholder farmers' incomes if yield losses – here defined as yield anomaly below the mean yield level according to Eq. 2 and Finger (Finger, 2013) – occur which are attributable to weather impacts. Our separation of maize yield loss factors shows dissimilar shares of weather-related (27%) and non-weather-related (73%) yield losses for Tanzania on average. Across districts, weather-related yield variability varies between 4% (in sub-humid south–east Tanzania) and 57% (in the semi-arid central and north–west, see Fig. 3). In total, the average and maximum weather-related yield losses are 0.11 and 0.41 t ha^{-1} and the non-weather-related yield losses are 0.34 and 1.70 t ha^{-1}, respectively. In line with the results of Lesk et al. (Lesk et al., 2016), this indicates that agronomic management and socio-economic factors have a substantially higher impact on maize yields in Tanzania (see also pre-analysis of significant non-weather-related yield effects in SI 2.3.1).

Our weather-attributable yield losses are directly usable to calculate claims for an area-based insurance scheme applied on Tanzanian district scale. The insurance claims are the product of the annual and district-specific weather-related yields losses (i), its corresponding maize acreage (ii), and the Tanzanian annual maize prices (iii). For the Tanzanian maize production, we calculate that the insurance of weather-related yield losses requires 71 million US\$ p.a. (23 US\$ ha^{-1}). In comparison, the insurance of the total yield losses would require 212 million US\$ p.a. (85 US\$ ha^{-1}). This means that 66% (141 million US\$ p.a.) of the loss costs for the insurance claims can be saved by indemnifying only weather-related yield losses calculated by our insurance approach (Fig. 3 bottom). Moreover, the trigger for paying claims would be rather below the arithmetic average (as in Fig. 3), because yield anomalies slightly below the arithmetic average are no actual losses and would be interpreted by farmers as "average yield". By choosing the 25%-percentile, the claim costs decrease to 29 million US\$ p.a. (10 US\$ ha^{-1}) and to 12 million US\$ p.a. (4 US\$ ha^{-1}) for the 10%-percentile, respectively. This reduces the premiums and by this, increases the affordability for smallholder farmers. The premium costs are determined by claim costs and transaction costs. The latter are not considered here, but typically range between 10 and 20% of the claim costs, i.e. the premium costs that the farmers would have to pay are dominated by the claim costs (Meze-Hausken et al., 2009; Shen and Odening, 2013). Since smallholder farmers often do not have the financial capacity to bear the insurance premiums, subsidization will be necessary. These subsidies can be disbursed from the government, like in the crop insurance system of the USA (Coble and Barnett, 2013), or from the Green Climate Fund, which supports climate adaptation projects, notably in SSA (Hochrainer-Stigler et al., 2014).

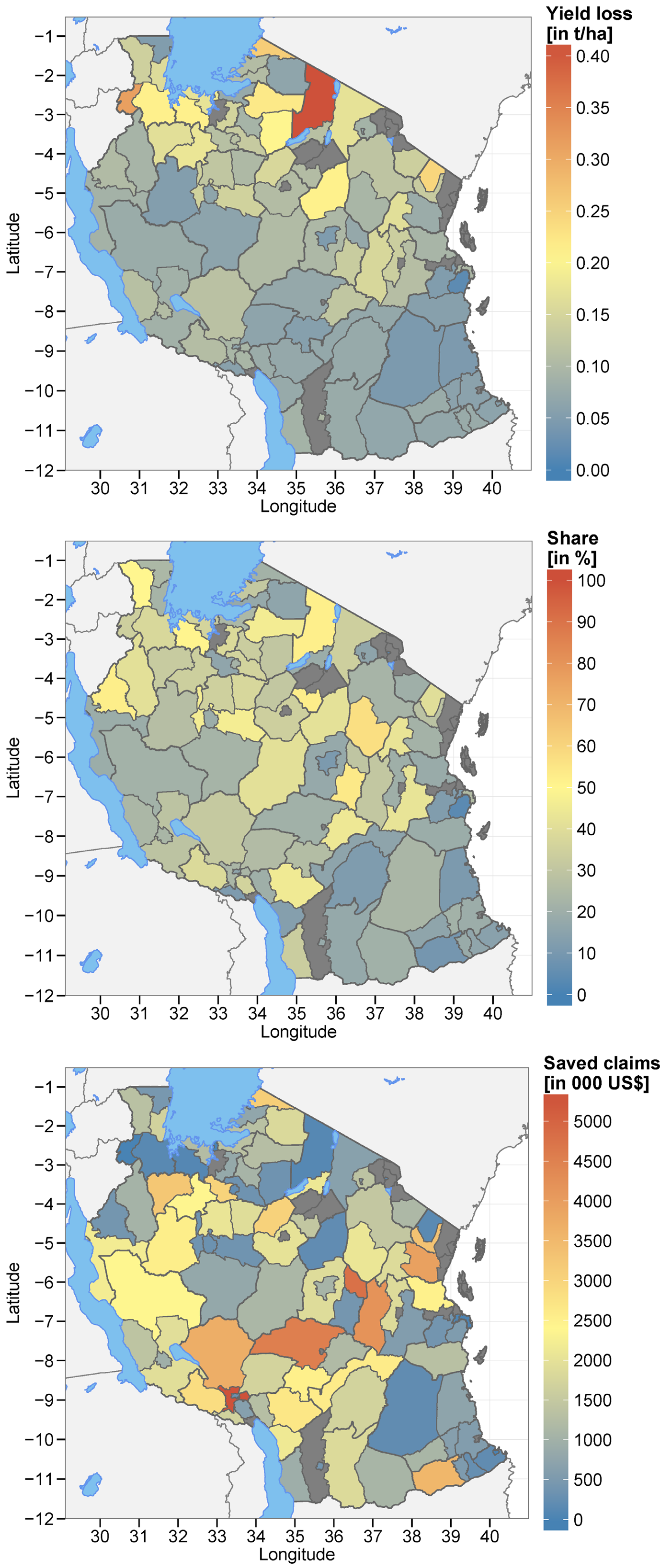


Fig. 3. Weather-related yield losses in t ha^{-1} p.a. (top), the share of weather-related yield losses in comparison to the total yield losses in % (middle), and saved costs for insurance claims in '000 US$ p.a. (bottom).

Our insurance index considers the uncertainty of the modeling approach. Since both crop model types still have limitations, we consider the model uncertainty for the insurance scheme. On the basis of weather-related yield variability, we calculate an insurance index, which is adjusted to the district-specific accuracy of the model approach (see *methods* for further information). Depending on the district scale model accuracy (R^2), we use weighted shares of modeled weather-related and observed yields for our insurance index (Eq. 3). Where the model is able to fully explain actual yield variability (by weather and non-weather-related impacts), the insurance index only uses the modeled weather-related yield variability of the process-based model (see Arusha, Kilimanjaro in Fig. 4). The share of observed yield variability increases in the index (for instance in Dodoma or Dar es Salaam) by decreasing the goodness of fit of our combined modeling approach. Due to the consideration of the model uncertainty, the claims increase to 141 million US$ p.a. (49 US$ ha^{-1}). Initially, this suggests it is less attractive for smallholder farmers to participate in the MAYIS solution. However, good and transparent coverage of the actual yield behavior will enhance the acceptance of the insurance scheme and thus improve the chance for a successful implementation of the insurance scheme (see SI S.3.3). In comparison to weather index insurances, the MAYIS has still lower premiums, because the non-weather-related perils are not indemnified.

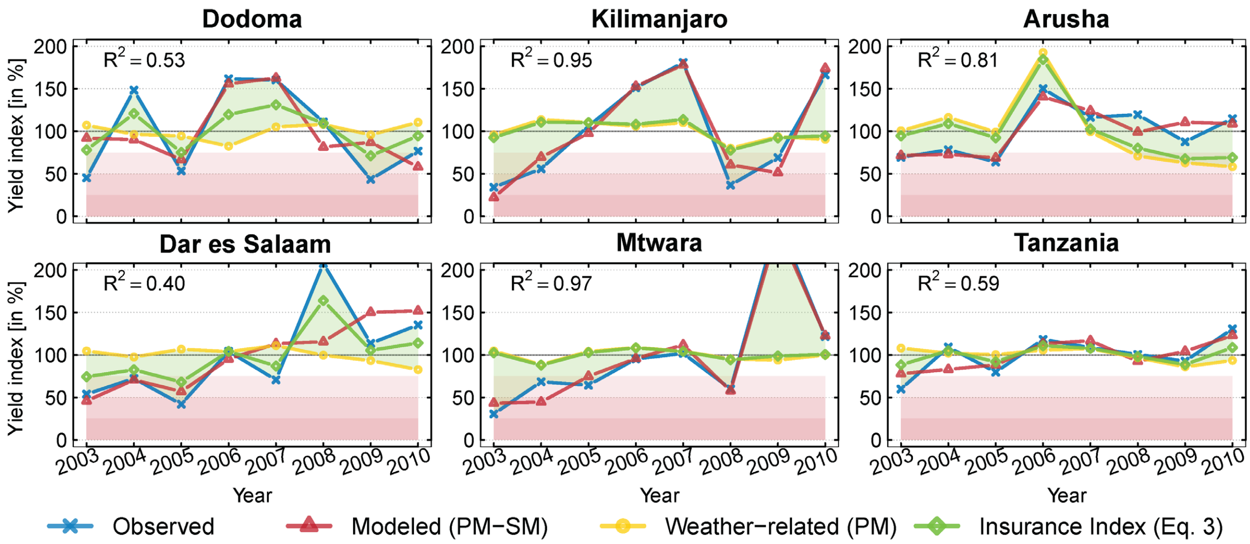


Fig. 4. Observed and modeled yields at regional scale for different agro-ecological zones and for entire Tanzania. The weather-related part is represented by the PM and the combination of weather and non-weather-related by the PM-SM modeling approach. The insurance index is calculated as R^2-weighted product of the observed and modeled yield variability. The R^2 is the goodness of fit for the modeled and the observed yields.

Our insurance index needs to be integrated in an implementation scheme. Although our study focuses on the accurate claim determination, we emphasize the relevance of a broader implementation scheme (see SI S.3 for further details). This should include, for instance, a legal framework, a concept to increase farmers' awareness of and trust in insurance schemes, and a mechanism to disburse claims to individual farmers. However, yield insurances are no panaceas. The chances for a successful implementation increase when such solutions are imbedded in (micro) credit agreements with local banks and coupled with education and agronomic management programs (Carter et al., 2016; Patt et al., 2010). Currently, Tanzanian farmers have no widespread access to crop insurance solutions as in many countries in SSA (Hochrainer-Stigler et al., 2014; Qureshi and Reinhard, 2014), but there are some promising activities. Among them are Agriculture and Climate Risk Enterprise (ACRE Africa) and African Risk Capacity (ARC). ACRE Africa has over 800,000 insured farmers in Kenya, Rwanda and Tanzania. ARC is designed to cover the country risk, but does not provide direct insurances for farmers. It aims to support the 30 member states (e.g., Ethiopia, Kenya, Malawi, and Zimbabwe in eastern Africa) to improve their capacities to cope with extreme weather events. These two examples already have an implementation scheme and an existing network of stakeholders with access to smallholder farmers. In regard of the loss determination, our approach could become a central component of an already existing implementation scheme, because it provides high accuracy to assess crop yield losses. This can contribute to build trust of the smallholder farmers in insurance solutions (Patt et al., 2009) and incentivize lenders offering credits for these farmers. An effective and accepted insurance solution can stabilize smallholder farmers' incomes and facilitate coping with changing weather patterns. The insurance claims will help to prevent farmers from losing or having to sell their livelihoods in years of extreme yield losses. In particular, they enable farmers to purchase food (after a yield loss) and agronomic inputs (in the following growing season). The insured incomes will give the farmers higher creditworthiness and thus, access to micro-credits for investments in production techniques, whose purchase is too risky without the insurance (Carter et al., 2016; Cole et al., 2013). As a result, this will enhance food security, indemnify livelihoods, and can have positive impacts on the health situation, migration reduction and even on the lives of smallholder farmers (Meze-Hausken et al., 2009; Wouterse, 2010).

6.4 Conclusions

The combination of the statistical and process-based crop modeling increases the accuracy of assessing actual yield variability. In our approach, we capture the plant-physiological yield development within the process-based model and large amounts of the remaining, unexplained yield variability by using a statistical model. The improvement in accuracy and robustness makes our approach suitable for crop production risk assessments on a district scale. Among other constraints for a successful and sustainable implementation, inadequate yield and yield impact information inhibit widespread implementation of index-based crop insurance schemes. We show that the suggested approach can contribute towards establishing a successful insurance scheme in Tanzania and other regions in SSA. This can reduce the vulnerability to severe yield losses for smallholder farmers and enhance farmers' ability to cope with climate change and altering weather patterns. Furthermore, the suggested area-based yield insurance scheme can contribute to long-term food security by incentivizing higher investments into agricultural production techniques.

6.5 Materials and methods

We apply a combined process-based (PM) and statistical (SM) modeling approach (PM-SM) to capture weather-attributable and non-weather-related yield variability. The PM captures influences on yield variability directly attributable to weather. The residual, non-weather-related yield variability of the process-based model is then modeled by a SM (see also SI Fig. S.3).

Process-based modeling of the weather-attributable yield variability

As PM we use the Soil and Water Integrated Model (SWIM). SWIM is an eco-hydrological model to capture river discharge, land use, and agricultural crop yield development (Krysanova et al., 2015, 2000). The crop module of SWIM is a modified approach of the Erosion Productivity Impact Calculator (EPIC) model (see also SI S.1.3 for further description). SWIM computes crop yields as a product of total above-ground biomass and the harvest index. Any divergence from the optimal growing conditions reduces biomass growth by stress factors within a minimum function. Considered stress factors are heat stress and water, nitrogen, and phosphorus scarcity. SWIM considers several agronomic management measures like fertilization, planting and harvest dates, and crop variety selection by maturity groups.

Statistical modeling of the non-weather-related yield variability

For our statistical model, we use a similar statistical approach to the approach used by Gornott and Wechsung (2016). The SM captures spatial and temporal heterogeneity in the residual yield variability of the PM. The SM estimates district-specific yield influences within a logarithmic function (Eq. 1). We use the statistical model with the residuals (ε_{it}) between the observed (y_{it}) and the process-based modeled yields (y_{it}^{PM}) as the endogenous variable and a vector of J exogenous variables (x_{jti}). The exogenous variables are maize acreage (in ha), paid subsidies on crop production (in US$), and urea application (in tons for entire Tanzania). Time-constant effects like land tenure security or market access (see SI S.1.4.3) are captured by the district-individual intercept (β_{0i}).

$$\varepsilon_{it} = \log \beta_{0i} + \sum_{j=1}^{J} \beta_{ji} \log\, x_{jit} + \log u_{it}, \qquad (1)$$

with β as parameters and u_{it} as error term for T years ($t = 1, \dots, T$) and N spatial units ($i = 1, \dots, N$).

Maize yield losses and insurance index

The mean weather and non-weather attributable yield loss (average yield anomaly below mean yield level) is calculated as semi-standard deviation (SSD_i^{below}, Eq. 2) for each district:

$$SSD_i^{below} = \sqrt{(T-1)^{-1} \sum_{t=1}^{T} \min\big((y_{it} - \bar{y}_i), \quad 0\big)^2}\,, \tag{2}$$

with $\bar{y}$ as arithmetic average yield across the T years.

The average indemnity claims are the product of SSD_{below}, maize acreage and maize price. In our case, the critical value for indemnity payments is the average yield. But other critical values, like the 25%-percentile and 10%-percentile, are also applied.

The maize yield insurance index (II) is calculated by Eq. 3. As maize yield insurance index (II), we use a weighted product of process-based modeled weather-related and observed yield variability. Depending on the accuracy of the combined model approach to explain the total yield variability, we weigh the share of modeled weather-related and total observed yield variability by the model R^2. Where the model is able to fully ($R^2 = 1$) explain total yield variability (by weather and non-weather-related impacts), only the weather-related modeled yield variability is used as the index. With decreasing R^2, the share of observed yield variability increases in the index (Eq. 3). The insurance index is normalized with the average yield and the factor 100.

$$II_{it} = 100\left(\left(\frac{y_{it}^{PM}}{\bar{y}_i^{PM}}\right)\mathrm{R}_i^2 + \left(\frac{y_{it}}{\bar{y}_i}\right)\left(1 - \mathrm{R}_i^2\right)\right) \tag{3}$$

To come to a claim payout, three steps are necessary: (1) Modeling of weather-attributable yield with (solely) weather data (2) calculation of the indemnity on bases of the historical weather-related yield distribution for each district, and (3) payout of the indemnity (see SI S.3.1.ii for details).

.

6.6 References

Affholder, F., Poeydebat, C., Corbeels, M., Scopel, E., Tittonell, P., 2013. The yield gap of major food crops in family agriculture in the tropics: Assessment and analysis through field surveys and modelling. F. Crop. Res. 143, 106–118.

Asseng, S. et al. 2013. Uncertainty in simulating wheat yields under climate change. Nat. Clim. Chang. 3, 827–832.

Bassu, S., Brisson, N., Durand, J.-L., Boote, K., Lizaso, J., Jones, J.W., Rosenzweig, C., Ruane, A.C., Adam, M., Baron, C., Basso, B., Biernath, C., Boogaard, H., Conijn, S., Corbeels, M., Deryng, D., De Sanctis, G., Gayler, S., Grassini, P., Hatfield, J., Hoek, S., Izaurralde, C., Jongschaap, R., Kemanian, A.R., Kersebaum, K.C., Kim, S.-H., Kumar, N.S., Makowski, D., Müller, C., Nendel, C., Priesack, E., Pravia, M.V., Sau, F., Shcherbak, I., Tao, F., Teixeira, E., Timlin, D., Waha, K., 2014. How do various maize crop models vary in their responses to climate change factors? Glob. Chang. Biol. 20, 2301–2320.

Benson, T., Kirama, S.L., Selejio, O., 2012. The supply of inorganic fertilizers to smallholder farmers in Tanzania evidence for fertilizer policy development. IFPRI Discuss. Pap. 1230, 1–48.

Blanc, E., 2012. The impact of climate change on crop yields in Sub-Saharan Africa. Am. J. Clim. Chang. 1, 1–13.

Carter, M.R., 2012. Designed for development impact: next generation approaches to index insurance for small farmers, in: Churchill, C., Reinhard, D. (Eds.), Protecting the Poor - Microinsurance Compendium Vol. II.

Carter, M.R., Cheng, L., Sarris, A., 2016. Where and how index insurance can boost the adoption of improved agricultural technologies. J. Dev. Econ. 118, 59–71.

Coble, K.H., Barnett, B.J., 2013. Why Do We Subsidize Crop Insurance? Am. J. Agric. Econ. 95, 498–504.

Cole, S., Giné, X., Tobacman, J., Topalova, P., Townsend, R., Vickery, J., 2013. Barriers to Household Risk Management: Evidence from India. Am. Econ. J. Appl. Econ. 5, 104–135.

Conradt, S., Finger, R., Bokusheva, R., 2015. Tailored to the extremes: Quantile regression for index-based insurance contract design. Agric. Econ. 46, 537–547.

Estes, L.D., Beukes, H., Bradley, B. a, Debats, S.R., Oppenheimer, M., Ruane, A.C., Schulze, R., Tadross, M., 2013. Projected climate impacts to South African maize and wheat production in 2055: a comparison of empirical and mechanistic modeling approaches. Glob. Chang. Biol. 19, 3762–3774.

Finger, R., 2013. Investigating the performance of different estimation techniques for crop yield data analysis in crop insurance applications. Agric. Econ. 44, 217–230.

Folberth, C., Gaiser, T., Abbaspour, K.C., Schulin, R., Yang, H., 2012. Regionalization of a large-scale crop growth model for sub-Saharan Africa: Model setup, evaluation, and estimation of maize yields. Agric. Ecosyst. Environ. 151, 21–33.

Gornott, C., Wechsung, F., 2016. Statistical regression models for assessing climate impacts on crop yields: A validation study for winter wheat and silage maize in Germany. Agric. For. Meteorol. 217, 89–100.

Herbold, J., 2014. New Approaches to Agricultural Insurance in Developing Economies, in: Köhn, D. (Ed.), Finance for Food: Towards New Agricultural and Rural Finance. Springer Berlin Heidelberg, p. 199–217.

Hochrainer-Stigler, S., Mechler, R., Pflug, G., Williges, K., 2014. Funding public adaptation to climate-related disasters. Estimates for a global fund. Glob. Environ. Chang. 25, 87–96.

Iizumi, T., Ramankutty, N., 2015. How do weather and climate influence cropping area and intensity? Glob. Food Sec. 4, 46–50.

IPCC, 2014. Africa, in: Niang, I., Ruppel, O.C., Abdrabo, M.A., Essel, A., Lennard, C., Padgham, J., Urquhart, P. (Eds.), Impacts, Adaptation and Vulnerability - Contributions of the Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. p. 1199–1265.

Jayne, T.S., Mather, D., Mason, N., Ricker-Gilbert, J., 2013. How do fertilizer subsidy programs affect total fertilizer use in sub-Saharan Africa? Crowding out, diversion, and benefit/cost assessments. Agric. Econ. 44, 687–703.

Krysanova, V., Hattermann, F., Huang, S., Hesse, C., Vetter, T., Liersch, S., Koch, H., Kundzewicz, Z.W., 2015. Modelling climate and land-use change impacts with SWIM: lessons learnt from multiple applications. Hydrol. Sci. J. 60, 606–635.

Krysanova, V., Wechsung, F., Arnold, J., Srinivasan, R., Williams, J., 2000. Soil and Water Integrated Model: User manual. Pik Rep. 69, 1–243.

Leblois, A., Quirion, P., 2013. Agricultural insurances based on meteorological indices: realizations, methods and research challenges. Meteorol. Appl. 20, 1–9.

Lesk, C., Rowhani, P., Ramankutty, N., 2016. Influence of extreme weather disasters on global crop production. Nature 529, 84–87.

Liu, B. et al., 2016. Similar estimates of temperature impacts on global wheat yield by three independent methods. Nat. Clim. Chang. 6, 1130–1136.

Lobell, D.B., Asseng, S., 2017. Comparing estimates of climate change impacts from process- based and statistical crop models. Environ. Res. Lett. 12, 1–12.

Lobell, D.B., Burke, M.B., 2010. On the use of statistical models to predict crop yield responses to climate

change. Agric. For. Meteorol. 150, 1443–1452.
Lobell, D.B., Ortiz-Monasterio, J.I., Asner, G.P., Matson, P. a., Naylor, R.L., Falcon, W.P., 2005. Analysis of wheat yield and climatic trends in Mexico. F. Crop. Res. 94, 250–256.
McIntosh, C., Sarris, A., Papadopoulos, F., 2013. Productivity, credit, risk, and the demand for weather index insurance in smallholder agriculture in Ethiopia. Agric. Econ. 44, 399–417.
Meze-Hausken, E., Patt, A., Fritz, S., 2009. Reducing climate risk for micro-insurance providers in Africa: A case study of Ethiopia. Glob. Environ. Chang. 19, 66–73.
Moore, N., Alagarswamy, G., Pijanowski, B., Thornton, P., Lofgren, B., Olson, J., Andresen, J., Yanda, P., Qi, J., 2012. East African food security as influenced by future climate change and land use change at local to regional scales. Clim. Change 110, 823–844.
Müller, C., Cramer, W., Hare, W.L., Lotze-Campen, H., 2011. Climate change risks for African agriculture. Proc. Natl. Acad. Sci. 108, 4313–4315.
Müller, C., Elliott, J., Chryssanthacopoulos, J., Arneth, A., Balkovic, J., Ciais, P., Deryng, D., Folberth, C., Glotter, M., Hoek, S., Iizumi, T., Izaurralde, R.C., Jones, C., Khabarov, N., Lawrence, P., Liu, W., Olin, S., Pugh, T.A.M., Ray, D., Reddy, A., Rosenzweig, C., Ruane, A.C., Sakurai, G., Schmid, E., Skalsky, R., Song, C.X., Wang, X., de Wit, A., Yang, H., 2016. Global Gridded Crop Model evaluation: benchmarking, skills, deficiencies and implications. Geosci. Model Dev. Discuss. 1–39.
Patt, A., Peterson, N., Carter, M., Velez, M., Hess, U., Suarez, P., 2009. Making index insurance attractive to farmers. Mitig. Adapt. Strateg. Glob. Chang. 14, 737–753.
Patt, A., Suarez, P., Hess, U., 2010. How do small-holder farmers understand insurance, and how much do they want it? Evidence from Africa. Glob. Environ. Chang. 20, 153–161.
Qureshi, Z., Reinhard, D., 2014. Microinsurance learning sessions Tanzania 2014 - New opportunities in a growing market. Report 1–33.
Ramirez-Villegas, J., Challinor, A., 2012. Assessing relevant climate data for agricultural applications. Agric. For. Meteorol. 161, 26–45.
Ray, D.K., Gerber, J.S., MacDonald, G.K., West, P.C., 2015. Climate variation explains a third of global crop yield variability. Nat. Commun. 6, 1–9.
Rosenzweig, C., Iglesias, A., Yang, X.B., Epstein, P.R., Chivian, E., 2001. Climate change and extreme weather events: Implications for food production, plant diseases, and pests. Glob. Chang. Hum. Heal. 2, 90–104.
Rötter, R.P., Carter, T.R., Olesen, J.E. and Porter, J.R., 2011. Crop-climate models need an overhaul. Nat. Clim. Chang., 1: 175–177.
Rowhani, P., Lobell, D.B., Linderman, M., Ramankutty, N., 2011. Climate variability and crop production in Tanzania. Agric. For. Meteorol. 151, 449–460.
Sánchez, P.A., 2010. Tripling crop yields in tropical Africa. Nat. Geosci. 3, 299–300.
Schauberger, B., Gornott, C., Wechsung, F., 2017. Global evaluation of a semiempirical model for yield anomalies and application to within-season yield forecasting. Glob. Chang. Biol.
Schlenker, W., Lobell, D.B., 2010. Robust negative impacts of climate change on African agriculture. Environ. Res. Lett. 5, 1–8.
Shen, Z., Odening, M., 2013. Coping with systemic risk in index-based crop insurance. Agric. Econ. 44, 1–13.
Surminski, S., Bouwer, L.M., Linnerooth-Bayer, J., 2016. How insurance can support climate resilience. Nat. Clim. Chang. 6, 333–334.
van der Velde, M., Folberth, C., Balkovič, J., Ciais, P., Fritz, S., Janssens, I.A., Obersteiner, M., See, L., Skalský, R., Xiong, W., Peñuelas, J., 2014. African crop yield reductions due to increasingly unbalanced nitrogen and phosphorus consumption. Glob. Chang. Biol. 20, 1278–1288.
Ward, P.S., Florax, R.J.G.M., Flores-Lagunes, A., 2014. Climate change and agricultural productivity in Sub-Saharan Africa: A spatial sample selection model. Eur. Rev. Agric. Econ. 41, 199–226.
Woodard, J.D., Garcia, P., 2008. Weather derivatives, spatial aggregation, and systemic risk: Implications for reinsurance hedging. J. Agric. Resour. Econ. 33, 34–51.
Wouterse, F., 2010. Migration and technical efficiency in cereal production: evidence from Burkina Faso. Agric. Econ. 41, 385–395.

7 General discussion

In the first article of this dissertation, we develop a statistical crop model and test a large set of potential agronomic management, policy and phenological divided weather variables for the German wheat production. The aim of this model is to decompose weather and non-weather yield influences on an aggregated level (i.e. federal states). In the second article, we develop a comparable statistical crop model approach applied on lower spatial scale (i.e. counties) on winter wheat and silage maize yields in Germany. The aim of this analysis is to test two further statistical regression methods to explain yield variability on a regional and aggregated level. This model is applied in the third article to investigate the performance for climate projections on regional and aggregated scale by analyzing the out-of-sample cross validation accuracy. Moreover, the impacts of weather influences within the vegetative and reproductive period as well as the non-weather influences on winter wheat and silage maize yields are shown. In the fourth article, we demonstrate that this approach is up-scalable to a global approach and to grain maize, spring wheat, and soybeans. This work closes with a model performance analysis on crop yield forecasts one or two moth prior to the scheduled harvest time. In the last article, we develop a combined statistical and process-based crop model approach to assess crop yield losses and decompose maize yield variability for index-insurance solutions in Tanzania. Finally, we discuss a framework for a potential insurance implementation scheme.

7.1 Decomposing weather- and management-related impacts on crop yields

One main finding of this dissertation is that the very low, but highly volatile maize yields in Tanzania are more sensitive to agronomic management than to weather variability. This is – at a first glance – somewhat surprising, but can be explained by the mostly favorable weather conditions in most regions of Tanzania. This is a very crucial outcome when looking at food security, because it demonstrates that there is a high potential to increase and stabilize crop yields through improved management. This gives farmers the opportunity to control large shares of their yield variability.

According to Liebig's law of the minimum, the crop yield is limited by the scarcest growing factor. In Tanzania, maize yields are very low (national average 1.3 t ha^{-1}), but have high inter-annual variability (standard deviation: ±0.9 t ha^{-1}), while the German (grain) maize yields attained 7.9 t ha^{-1} in the same period (2003–2010). The low yield level in combination with the high variability indicates that the Tanzanian yields are highly sensitive to either weather or agronomic management. Yield limitations caused by poor soil quality are also common in SSA, but would not explain the high inter-annual variability. Because of the mostly favorable weather conditions in Tanzania (Thornton et al., 2010; van Ittersum et al., 2013), management factors should be hypothetically the main driver of yield variability in this region. Hence, the yield influence of weather should be smaller than agronomic management-related influence. In line with this hypothesis, the management-driven statistical model is able to robustly resolve 73% of the temporal and spatial yield variability by using agronomic management and socio-economic variables only. Moreover, we find in our analyses, that neither the process-based model (with constant agronomic management) nor the purely weather-driven statistical model can reproduce the high maize yield variability in Tanzania. Thus, we conclude that most of the maize yield variability (73%) is attributable to agronomic management and socio-economy and only 27% is attributable to weather impacts. The weather influence is even less than in the results of Rowhani et al. (2011), who also use a statistical crop model, which explains 34 to 41% (re-calculated form the adjusted R^2 of 0.324 and 0.395, $N = 19$, $T = 14$, $K = 7$) of Tanzanian maize yield variability by weather factors at regional scale. However, our analysis is conducted one administrate level lower at district scale ($N = 116$). Moreover, our weather-related yield variability is calculated with the process-based model, which does not control for collinear weather-triggered yield impacts of pests and diseases as statistical models. Hence, the weather-attributable yield variability of our analysis might be slightly higher if we also took into account the yield variability explained by our statistical model, which controls for indirect weather-triggered yield impacts. However, this statistical model shows no robust results.

For our aggregated analysis (Albers et al., 2017) , we decompose the influence of weather, agricultural management and agricultural policy on wheat yields at German federal state level ($N = 12$). In this analysis, we find that agronomic management explains 49% of the actual wheat yield variability, while weather explains 43%. At the German county level ($N = 289$), Conradt et al. (2016) show that winter wheat and silage maize yield variability is mainly attributable to weather influences. Due to the lower aggregation level of the input data, the statistical model of Conradt et al. (2016) and Gornott and Wechsung (2016, 2015) are able to capture county-specific weather impacts. These two results show that the chosen spatial scale makes a great difference. Also in Tanzania, the weather-attributable maize yield variability is higher on an aggregated level (Rowhani et al., 2011) than on the scale of one administrate level lower, meaning at the district scale (results of our analysis). This finding is in line with our results at German federal states level (Albers et al., 2017) in comparison to the county level (Conradt et al., 2016; Gornott and Wechsung, 2016). Thus, we conclude that the weather-related yield variability is higher at aggregated level. Furthermore, at comparable spatial scale, weather-related yield variability is higher in Germany than in Tanzania. This can be explained by the low Tanzanian maize yields, which are mostly limited by an insufficient agronomic management. Accordingly, large shares of the actual yield losses could be prevented by improving agronomic management (Thornton et al., 2010; van Dijk et al., 2017). For Germany, this demonstrates that the agronomic management is constantly on high and well-organized level, which leaves less space for improvements.

7.2 Disaggregation of the growing season and determination of sub-periods

In most of the statistical crop models (Moore and Lobell, 2014; You et al., 2009), weather data is aggregated over the entire growing period (see Albers et al., 2017: SI Tab. S.2 for a detailed literature list). For statistical models using growing periods divided in sub-periods, weather data is mostly aggregated by calendar months (Heimfarth et al., 2012; Lobell et al., 2005; Miao et al., 2016). Only a few studies use observed phenological development stages to further break down the growing season (Butler and Huybers, 2015; Dixon et al., 1994). Analogues to the literature, we also consider different temporal aggregation levels for the weather variables in our statistical crop models for Germany and Tanzania. In Gornott and Wechsung (2016, 2015), our weather variables account for the weather-related yield impacts during the vegetative and reproductive growing period. For that, we use the average growing periods for winter wheat and silage maize for entire Germany calculated with the data provided by the German Weather Service (DWD). In the global statistical modeling approach (Schauberger et al., 2017), we do the same with a dataset providing global planting and harvesting dates (MIRCA2000). Since this dataset contains only planting and harvesting dates, we divide the cumulated growing degree days within the region-specific growing period by half of these cumulated growing degree days. In the approach focusing on Tanzania, we consider district-specific planting and harvesting dates (FAO Crop Calendar). Both latter approaches (the global approach and the approach for Tanzania) use planting and harvesting dates, which are constant over time, but vary across space. In our aggregated statistical approach for Germany (Albers et al., 2017), we use four phenological development periods, which vary across space and time. This allows controlling the different impacts of weather on crop yields in the main phenological stages, which is not possible in the approach of Gornott and Wechsung (2016, 2015). The aggregation over the entire reproductive growing period (as in Gornott and Wechsung, 2016, 2015) does not account for different crop requirements in the stages from the heading to anthesis and ripening to harvesting as it is discussed by Rötter and Geijn (1999). For instance, this analysis on German county level shows only a weak negative precipitation effect in the reproductive period (Gornott and Wechsung, 2016). After decomposing into phenological relevant periods, we find significant yield sensitivity to heat in the heading to anthesis and a negative impact of precipitation between ripening and harvesting (Albers et al., 2017).

When looking at the different statistical models for Germany, the aggregated PDM approach (Albers et al., 2017), despite using phenological development, achieves a similar goodness of fit ($R^2 = 0.83$) than our county-scale (Gornott and Wechsung, 2016) STSM approach ($R^2 = 0.86$). Also other applications derive similar goodness fits by models using a division in phenological phases and those using calendar months (Dixon et al., 1994). However, our county-scale PDM approach (Gornott and Wechsung, 2016) has a significantly lower goodness of fit ($R^2 = 0.69$) than the aggregated PDM approach (Albers et al., 2017). This difference in goodness of fit of both panel data models could also be a result of the aggregation level or variable selection. Although it is preferable to use phenological information for the variable division, mostly this data is not available, notably not for countries like Tanzania. Finally, as forecasts of phenological dates are very uncertain (Ma et al., 2012), weather variables divided by phenological data are not applicable for yield forecasts and climate change projections. Notwithstanding, high estimation and validation goodness fits are also possible without using this data.

7.3 Crop yield projections

Medium- and long-term climate change projections are important for developing strategies to support farmers to cope with climate change impacts on crop yields. These climate change projections can be conducted with both statistical and process-based crop models (Liu et al., 2016; Müller et al., 2011). Process-based models are designed to project future climate impacts on yields by considering changes in agronomic management (like irrigation), atmospheric CO_2 concentration, and extreme weather conditions (Elliott et al., 2014; Rosenzweig et al., 2014). Since process-based models frequently face the problem of biased climate data, these models , however, require complex bias correction measures to carry out sound climate change projections on crop yields (Hawkins et al., 2013; Lobell, 2013). Statistical crop models project yields by coupling their parameters with climate datasets of future periods, but can only partly capture long-term yield trends due to increasing temperatures or atmospheric CO_2 concentrations (Rötter et al., 2011). Therefore, approaches are needed which do not base on the absolute values and overcome an explicit trend modeling. In comparison to most other statistical models (see e.g. Butler and Huybers, 2015; Lobell et al., 2014; Ward et al., 2014), our statistical models estimate relative yield changes instead of absolute yield levels. This transformation eliminates any potential linear bias and trend in the exogenous and endogenous variables. Hence, an explicit bias correction and modeling of crop yield trends is redundant. Nevertheless, the application of statistical models also has limitations, which diminish their suitability for climate change projection. These limitations are, for instance, a limited ability of statistical models to account for extremes and non-linear relationship between yields and their influencing factors beyond the observed range. To overcome this issue, a quantile regression can enhance the model accuracy, notably with the aim to increase the capacity to capture extremes (Conradt et al., 2015).

For the case of Tanzania, we use a process-based crop model to explain the weather-related yield variability. This model reproduces accurately yield levels and spatial yield variability, but insufficiently the inter-annual yield variability. However, a low reproduction accuracy of observed inter-annual yield variability is no indicator that the projection to future conditions also has a low accuracy (Müller, 2011; Müller et al., 2016). Because it could be that higher model fits are achieved by a thoroughly conducted calibration on the historical conditions, but this calibration might not be valid for future conditions. Moreover, the model may lack processes which gain in importance in the future (e.g. sensitivity to heat stress). Process-based models are applicable for a large range of environmental conditions, because of their bio-physical organization. Notwithstanding, process-based crop models also have limitations in regard to projecting yield impacts in future periods. These limitations are, for instance, the lack of high-quality input and reference data such as growing season dates or information on fertilizer applications (Müller et al., 2016), but also the quality of soil data contributes to uncertain yield assessments (Folberth et al., 2016). Moreover, in regions with low weather station density, fragmented and imprecise weather data contributes to the assessment uncertainty (Van Wart et al., 2013). Furthermore, the parameters in many process-based models (including the EPIC crop module in SWIM) are mostly derived from observations before the mid-1980 and thus, do not include past (last three decades) and recent developments in crop breeding (Rötter et al., 2011). However, at least for Tanzania and other regions in SSA, seed varieties and cropping practices did not substantially change in past years (McClung, 2014; Westengen et al., 2014). In addition, processes – which are not imbedded in the model, but also influence crop yields (like intercropping or tillage practices) – can influence the model results (Snapp et al., 2014). Although process-based and statistical models base on completely different approaches, both model types calculate similar yield changes attributable to climate change (Liu et al., 2016; Lobell and Asseng, 2017) and thus, contribute valuable projections for the agricultural sector.

7.4 Models' ability to forecast crop yields

Yield changes can have a leverage effect on food security namely that small yield losses amplify food insecurity (West et al., 2014). Short-term forecast instruments and early warning systems can support the handling of upcoming food shortages, which might lead to food insecurity. Based on these short-term crop yield forecasts, farmers can – if still possible – adjust their agronomic management and invest in strategies to cope with production risks like crop insurances or future contracts of agricultural commodities (Chipanshi et al., 2015; Qian et al., 2009; Stone and Meinke, 2005; Woodard and Garcia, 2008). In our global approach (Schauberger et al., 2017), we expand and apply short-term yield forecasts of up to two months prior to harvest and show that our approach is (at least in some world regions) suitable for yield forecasts. For these forecasts, high accuracy in assessing inter-annual yield variability is indispensable. In Gornott and Wechsung (2016, 2015), we show that statistical models are able to satisfactorily reproduce temporal and spatial yield variability within an out-of-sample cross-validation as it would be required for yield forecasts. In a consecutive study, Conradt et al. (2016) show that the validation performance can be further improved by slightly augmented exogenous variables and a cluster analysis. The latter is used to restructure aggregation units from administrative boundaries (federal states) to crop-specific agro-ecological zones. Moreover, the individual parametrization of our separate time series model (STSM) can also capture extreme yield anomalies, which occur only in single regions, with high accuracy (as shown by Gornott and Wechsung, 2015). This and the out-of-sample robustness of this statistical approach demonstrate the models' capacity to reproduce and project unknown yields with observed weather information of the growing season. For yield forecasts, this observed weather information is, however, not available in advance and thus, the approach would rely on weather forecast data. This consideration of weather forecast would add a further source of uncertainty, because weather forecasts are only reliable for one to four weeks in advance, but get very uncertain when it comes to projecting more than one month in advance. To reduce the uncertainty, only forecasts up to one month in advance should be used (Kusunose and Mahmood, 2016; Lee et al., 2013). Notwithstanding, the availability of reliable weather forecast data and the model capacity to bridge gaps of unavailable weather data (as shown in Schauberger et al., 2017) demonstrate the potential for yield forecasts by our approach. Such statistical yield forecast instruments can be used to support farmers by adjusting their agronomic management and help them to deal with production shortages.

7.5 Accuracy, acceptance and affordability of insurance solutions

A high accuracy of the insurance index (which is used as trigger for claim payouts) is crucial for the implementation and acceptance of an area-based yield insurance scheme. Thus, the selection of a proper index is important for the successes of the insurance scheme (Conradt et al., 2015; Leblois and Quirion, 2013). Weather index insurances (based on e.g. precipitation indexes) are easy to understand (McIntosh et al., 2013; Sarris, 2013), but often only poorly correlated with the actual crop yields. To determine yield losses attributable to weather-related impacts, cost and data efficient solutions are needed, which do not rely on assessments of claim adjusters. While process-based models can be applied without observed yield data, but require detailed information on management, soil and weather; statistical models can be applied with aggregated and incomplete management and weather information, but require observed yield data (at least for an estimation). Because of this, new approaches are desirable, which can deal with limited yield, management and weather information and achieve high assessment accuracy at the same time. In particular in regions with limited data availability like SSA, the assessment accuracy is often insufficient (Bassu et al., 2014; Grassini et al., 2015). For the case of Tanzania, we develop and test a combined statistical and process-based crop model approach. The strength of this approach is to deal with limited data availability and thus, it is able to overcome (at least partly) the issue of unavailable high-resolution management and socio-economic datasets. Moreover, our approach would allow incorporating weather-related yield data determined by other process-based crop models. The increased availability of this yield data calculated by process-based models – notably due to the Agricultural Model Intercomparison and Improvement Project (AgMIP) – further offers the opportunity to consider modeled yields with high accuracy for respective regions around the word (Asseng et al., 2013; Liu et al., 2016; Müller et al., 2016).

8 Conclusion

In this dissertation, we develop and test statistical crop model applications – which control for and decompose weather and agronomic management-related yield impacts – for the case of Germany and a global approach. By using this statistical methods and decomposing approach, we develop a new crop model approach by combing process-based and statistical crop models for the case of Tanzania. For this approach, we use a process-based crop model to capture the weather-related yield variability and a consecutive statistical model to cover the non-weather-related yield variability. Within this dissertation, we show that all applied statistical models as well as the combined modeling approach achieve robust results in reproducing and projecting crop yields for the different case study regions. This demonstrates that the applied crop models are suitable to assess yield variability in both temperate, intensively managed and tropical, extensively managed agricultural systems. Across the case study regions, one main finding is that the yields in Tanzania are more sensitive to agronomic management than to weather variability. In contrast to that the weather impact seems to be the main driver of yield variability in Germany at the county level. These findings demonstrate a high potential to increase and stabilize low crop yields in Tanzania through improved management and thus, allow farmers to enhance their food security situation by changing management practices. Moreover, we find that the statistical crop model applications are suitable for climate change impact analyses, in-season yield forecasts, and transferable to other crops and regions. Finally, we demonstrate that our crop model assessments allow applications concerning risk management (e.g. support investment and management decisions) and risk transfer (e.g. insurance solutions). These findings are helpful for the implementation of insurance solutions to stabilize smallholder farmers' incomes and enhance their food security situation.

9 Outlook

9.1 Increasing yield assessment accuracy and spatial resolution

In addition to the work developed in this dissertation, further improvements and adjustments of the applied crop models as well as further applications are thinkable and to some extend already in preparation. Relatively easy to implement are improvements of both model types by increasing the quality of the input data. For instance, statistical models might achieve higher accuracy by using augmented and more detailed data (on e.g. pests and diseases) and an advanced pre-processing of the variables (e.g. developing indicators for water or nutrient deficiency). This might allow considering more complex relationships between yield and yield influencing factors and thus, deliver a more precise knowledge of yield influencing factors. For the process-based models, it would be beneficial to expand the crop yield responses (where model account for) by imbedding further sub-processes of e.g. plant diseases spreading patterns and certain agronomic management practices. To do so, a selection of the relevant, but still unconsidered yield influences is required. Our combined approach might be helpful for selecting these relevant and so far in the process-based model unconsidered yield influences from the statistical model results. In a further step, these selected yield impacts could be formulated in sub-processes and imbedded in the process-based models. This can give way to better tailored policies and extension services for farmers and other stakeholders along the food value chain.

Further research should focus on increasing yield assessment accuracy and spatial resolution thereof. This can be conducted by using more detailed datasets like household surveys and crop field trials. However, it is likely that this data is not available for a wider scale, notably not for many regions in SSA. Thus, other data sources (stemming from e.g. precision farming data) and innovative approaches (which integrate this data) are needed. For instance, in many world regions, farming machinery meanwhile uses measurement instruments to document yields and agronomic management at sub-area (intra-field) level for internal farm and crop management strategies. For this purpose, combine harvesters quantify sub-area-specific crop yields and its anomalies, while fertilizer spreaders and crop protection sprayers document respective agronomic management measures. The latter is also used for documentation duties to the EU to receive CAP subsidies and other regulators demanding data on e.g. fertilizer application and plant protection sprayings. Such data sources would provide enormous amounts of additional and very detailed data. However, so far, this data are not publically available and mostly collected in high-input agricultural countries like Germany or the USA, but only limitedly in SSA. In addition to this data source, large amounts of social media platforms' data (e.g. Twitter) is mainly publically available around the world including SSA. To use this data for crop models, new and innovative approaches for data-mining and analyzing are required. For instance, tools like artificial intelligence (e.g. neuronal networks, machine learning) could be integrated in the existing model approaches to enable crop models to automatically incorporate new relationships of crop yield influences.

Moreover, it is possible to expand the combined process-based and statistical crop modeling approach with the help of remote sensing data to assess yield losses at field and local level. As in our modeling approach for Tanzania, the process-based model would capture influences on yield variability directly attributable to weather. The residual yield variability would be modeled by a consecutive statistical model to control for yield influences of agronomic management, socio-economic and indirect weather-triggered impacts (like plant health). As further information, the approaches should use vegetation indexes of remote sensing data (e.g. Leaf Area Index – LAI, Normalized Difference Vegetation Index – NDVI). For instance, the LAI data could be used to calibrate the process-based model to provide field level yields. The field level process-based model results could be supplemented by statistical model assessments using weather data to quantify influences of weather triggered plant health effects and control for agronomic management effects. Remote sensing data has a worldwide and constant coverage and is already used for precision farming and crop management in the agricultural sector. In the recent years, the quality, resolution and availability of remote sensing images and processing tools has substantial increased (in particular with the launched Sentinel data of the EU Copernicus program).

New satellite sensors provide high spatial, temporal and spectral (e.g. red-edge) resolution of optical data. Despite the promising developments in such technologies, this data are so far not used in crop model approaches although the improved data quality would allow integrating this information in the crop models. The incorporating of remote sensing data should increase the crop models' accuracy, spatial resolution, and availability of field level yield assessments immediately after the occurrence of a crop failure. As far as these assessments meet these specific requirements, these yield assessments are of high interest for insurances companies.

9.2 Possible application for the modeling approach

Generally, there is a high demand for short-term yield forecasts to predict food shortages and crop failures as indicator for food insecurity. By using standardized forecast weather data and the above discussed model improvements, it might be possible to run such yield forecasts automatically to provide weekly or monthly updates. Such forecasts would not contain expert-based yield assessments as still needed for other forecast systems (like MARS). Moreover, if such forecasts were available on a global scale, it would be possible to link them with commodity price models (e.g. Schewe et al., 2017) to gain information on upcoming price changes. These price models would use global crop yield and respective acreage information as supply data. In combination with demand and storage information, such models can predict price changes. The linkage between crop and price models might help to avert weather risks in the crop production and support farmers by choosing adequate input intensity for agronomic management. Since crop production and commodity prices affect food availability and access to food, this can also contribute to increase local food security (Gilbert et al., 2017; Wheeler and von Braun, 2013).

Weather induced food insecurity can amplify and cause health problems like child mortality or vector and water borne diseases like malaria and cholera (Phalkey et al., 2015; Stoltzfus et al., 2014) and thus, further increase chronic undernourishment, hidden hunger, and poverty. In SSA, crop production losses or failures are associated with diminished child nutrition and survival. In particular in rural areas, smallholder farmers are often unable to provide sufficient food (in terms of calories and nutritional values) for their family. The results of the here developed crop models can be used to assess the weather impacts on crop yields and its consequences for food availability. This information can be further used to quantify the impact of low crop yields and food availability shortages on child stunting and mortality with specific health models (e.g. Belesova et al., 2017) under the current and future climate conditions.

Moreover, food insecurity commonly induces household members to migrate. Knowledge about the food security status and the reasons for migration can help to understand the complex relationship between climate (and climate change) and unwanted migration. Often, migration is used as an adaptation strategy to generate (additional) income, because of the lack of other opportunities to adapt to weather extremes and food insecurity. This raises the question, if the migration pressure can be diminished by more efficient adaptation strategies and if so, what are suitable adaptation strategies. As one example, the implementation of crop insurance solutions is discussed as relevant adaptation strategy in the context of an increasing pressure caused by permanent, international and seasonal, domestic migration (Wouterse, 2010). The availability of crop insurance solutions can contribute to decrease unwanted migration and make it possible for people to stay in rural regions (notwithstanding the continued relevance of migration as an adaptation strategy).

10 References

Albers, H., Gornott, C., Hüttel, S., 2017. How do inputs and weather drive wheat yield volatility? The example of Germany. Food Policy 70, 50–61.

Asseng, S., Ewert, F., Rosenzweig, C., Jones, J.W., Hatfield, J.L., Ruane, A.C., Boote, K.J., Thorburn, P.J., Rötter, R.P., Cammarano, D., Brisson, N., Basso, B., Martre, P., Aggarwal, P.K., Angulo, C., Bertuzzi, P., Biernath, C., Challinor, A.J., Doltra, J., Gayler, S., Goldberg, R., Grant, R., Heng, L., Hooker, J., Hunt, L.A., Ingwersen, J., Izaurralde, R.C., Kersebaum, K.C., Müller, C., Naresh Kumar, S., Nendel, C., O'Leary, G., Olesen, J.E., Osborne, T.M., Palosuo, T., Priesack, E., Ripoche, D., Semenov, M.A., Shcherbak, I., Steduto, P., Stöckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Travasso, M., Waha, K., Wallach, D., White, J.W., Williams, J.R., Wolf, J., 2013. Uncertainty in simulating wheat yields under climate change. Nat. Clim. Chang. 3, 827–832.

Bassu, S., Brisson, N., Durand, J.-L., Boote, K., Lizaso, J., Jones, J.W., Rosenzweig, C., Ruane, A.C., Adam, M., Baron, C., Basso, B., Biernath, C., Boogaard, H., Conijn, S., Corbeels, M., Deryng, D., De Sanctis, G., Gayler, S., Grassini, P., Hatfield, J., Hoek, S., Izaurralde, C., Jongschaap, R., Kemanian, A.R., Kersebaum, K.C., Kim, S.-H., Kumar, N.S., Makowski, D., Müller, C., Nendel, C., Priesack, E., Pravia, M.V., Sau, F., Shcherbak, I., Tao, F., Teixeira, E., Timlin, D., Waha, K., 2014. How do various maize crop models vary in their responses to climate change factors? Glob. Chang. Biol. 20, 2301–2320.

Belesova, K., Gasparrini, A., Sié, A., Sauerborn, R., Wilkinson, P., 2017. Household cereal crop harvest and children's nutritional status in rural Burkina Faso. Environ. Heal. 16, 65.

Berg, A., Quirion, P., Sultan, B., 2009. Weather-Index Drought Insurance in Burkina-Faso: Assessment of Its Potential Interest to Farmers. Weather. Clim. Soc. 1, 71–84.

Blanc, E., 2012. The impact of climate change on crop yields in Sub-Saharan Africa. Am. J. Clim. Chang. 1, 1–13.

Butler, E.E., Huybers, P., 2015. Variations in the sensitivity of US maize yield to extreme temperatures by region and growth phase. Environ. Res. Lett. 10, 34009.

Butler, E.E., Huybers, P., 2012. Adaptation of US maize to temperature variations. Nat. Clim. Chang. 3, 68–72.

Challinor, A., Wheeler, T., Garforth, C., Craufurd, P., Kassam, A., 2007. Assessing the vulnerability of food crop systems in Africa to climate change. Clim. Change 83, 381–399.

Chipanshi, A., Zhang, Y., Kouadio, L., Newlands, N., Davidson, A., Hill, H., Warren, R., Qian, B., Daneshfar, B., Bedard, F., Reichert, G., 2015. Evaluation of the Integrated Canadian Crop Yield Forecaster (ICCYF) model for in-season prediction of crop yield across the Canadian agricultural landscape. Agric. For. Meteorol. 206, 137–150.

Christiaensen, L., 2017. Agriculture in Africa – Telling myths from facts: A synthesis. Food Policy 67, 1–11.

Cole, S., Giné, X., Tobacman, J., Topalova, P., Townsend, R., Vickery, J., 2013. Barriers to Household Risk Management: Evidence from India. Am. Econ. J. Appl. Econ. 5, 104–135.

Conradt, S., Finger, R., Bokusheva, R., 2015. Tailored to the extremes: Quantile regression for index-based insurance contract design. Agric. Econ. 46, 537–547.

Conradt, T., Gornott, C., Wechsung, F., 2016. Extending and improving regionalized winter wheat and silage maize yield regression models for Germany: Enhancing the predictive skill by panel definition through cluster analysis. Agric. For. Meteorol. 216, 68–81.

Dixon, B.L., Hollinger, S.E., Garcia, P., 1994. Estimating Corn Yield Response Models to Predict Impacts of Climate Change. J. ofAgricultural Resour. Econ. 19, 58–68.

Elliott, J., Deryng, D., Müller, C., Frieler, K., Konzmann, M., Gerten, D., Glotter, M., Flörke, M., Wada, Y., Best, N., Eisner, S., Fekete, B.M., Folberth, C., Foster, I., Gosling, S.N., Haddeland, I., Khabarov, N., Ludwig, F., Masaki, Y., Olin, S., Rosenzweig, C., Ruane, A.C., Satoh, Y., Schmid, E., Stacke, T., Tang, Q., Wisser, D., 2014. Constraints and potentials of future irrigation water availability on agricultural production under climate change. Proc. Natl. Acad. Sci. 111, 3239–3244.

Estes, L.D., Beukes, H., Bradley, B. a, Debats, S.R., Oppenheimer, M., Ruane, A.C., Schulze, R., Tadross, M., 2013. Projected climate impacts to South African maize and wheat production in 2055: a comparison of empirical and mechanistic modeling approaches. Glob. Chang. Biol. 19, 3762–3774.

Ewert, F., Rötter, R.P., Bindi, M., Webber, H., Trnka, M., Kersebaum, K.C., Olesen, J.E., van Ittersum, M.K., Janssen, S., Rivington, M., Semenov, M.A., Wallach, D., Porter, J.R., Stewart, D., Verhagen, J., Gaiser, T., Palosuo, T., Tao, F., Nendel, C., Roggero, P.P., Bartošová, L., Asseng, S., 2015. Crop modelling for integrated assessment of risk to food production from climate change. Environ. Model. Softw. 72, 287–303.

Ewert, F., van Ittersum, M.K., Heckelei, T., Therond, O., Bezlepkina, I., Andersen, E., 2011. Scale changes and model linking methods for integrated assessment of agri-environmental systems. Agric. Ecosyst. Environ. 142, 6–17.

FAO Country STAT, 2017. Tanzania. URL https://countrystat.org/home.aspx?c=TZA

FAO Stat, 2013. FAO Stat database collections: National maize yields (2003-2010) . URL http://faostat3.fao.org/

Finger, R., 2013. Investigating the performance of different estimation techniques for crop yield data analysis in

crop insurance applications. Agric. Econ. 44, 217–230.
Fischer, G., Shah, M., 2010. Farmland Investments and Food Security. World Bank Part 2: St.
Fishman, R., 2016. More uneven distributions overturn benefits of higher precipitation for crop yields. Environ. Res. Lett. 11, 1–7.
Folberth, C., Gaiser, T., Abbaspour, K.C., Schulin, R., Yang, H., 2012. Regionalization of a large-scale crop growth model for sub-Saharan Africa: Model setup, evaluation, and estimation of maize yields. Agric. Ecosyst. Environ. 151, 21–33.
Folberth, C., Skalský, R., Moltchanova, E., Balkovič, J., Azevedo, L.B., Obersteiner, M., van der Velde, M., 2016. Uncertainty in soil data can outweigh climate impact signals in global crop yield simulations. Nat. Commun. 7, 11872.
Foley, J. a, Ramankutty, N., Brauman, K. a, Cassidy, E.S., Gerber, J.S., Johnston, M., Mueller, N.D., O'Connell, C., Ray, D.K., West, P.C., Balzer, C., Bennett, E.M., Carpenter, S.R., Hill, J., Monfreda, C., Polasky, S., Rockström, J., Sheehan, J., Siebert, S., Tilman, D., Zaks, D.P.M., 2011. Solutions for a cultivated planet. Nature 478, 337–42.
German Federal Statistical Office, 2017. Landwirtschaftliche Betriebe . URL https://www.destatis.de/
Gilbert, C.L., Christiaensen, L., Kaminski, J., 2017. Food price seasonality in Africa: Measurement and extent. Food Policy 67, 119–132.
Gohin, A., 2006. Assessing CAP Reform: Sensitivity of Modelling Decoupled Policies. J. Agric. Econ. 57, 415–440.
Gornott, C., Wechsung, F., 2016. Statistical regression models for assessing climate impacts on crop yields: A validation study for winter wheat and silage maize in Germany. Agric. For. Meteorol. 217, 89–100.
Gornott, C., Wechsung, F., 2015. Level normalized modeling approach of yield volatility for winter wheat and silage maize on different scales within Germany | Niveauneutrale modellierung der ertragsvolatilität von winterweizen und silomais auf mehreren räumlichen ebenen in Deutschland. J. fur Kult. 67.
Grassini, P., Eskridge, K.M., Cassman, K.G., 2013. Distinguishing between yield advances and yield plateaus in historical crop production trends. Nat. Commun. 4, 2918.
Grassini, P., van Bussel, L.G.J., Van Wart, J., Wolf, J., Claessens, L., Yang, H., Boogaard, H., de Groot, H., van Ittersum, M.K., Cassman, K.G., 2015. How good is good enough? Data requirements for reliable crop yield simulations and yield-gap analysis. F. Crop. Res. 177, 49–63.
Grassini, P., Yang, H., Cassman, K.G., 2009. Limits to maize productivity in Western Corn-Belt: A simulation analysis for fully irrigated and rainfed conditions. Agric. For. Meteorol. 149, 1254–1265.
Hawkins, E., Osborne, T.M., Ho, C.K., Challinor, A.J., 2013. Calibration and bias correction of climate projections for crop modelling: An idealised case study over Europe. Agric. For. Meteorol. 170, 19–31.
Heimfarth, L.E., Finger, R., Musshoff, O., 2012. Hedging weather risk on aggregated and individual farm-level. Agric. Financ. Rev. 72, 471–487.
Hill, R.V., Hoddinott, J., Kumar, N., 2013. Adoption of weather-index insurance: learning from willingness to pay among a panel of households in rural Ethiopia. Agric. Econ. 44, 385–398.
Iizumi, T., Ramankutty, N., 2015. How do weather and climate influence cropping area and intensity? Glob. Food Sec. 4, 46–50.
Iizumi, T., Sakuma, H., Yokozawa, M., Luo, J.-J., Challinor, A.J., Brown, M.E., Sakurai, G., Yamagata, T., 2013. Prediction of seasonal climate-induced variations in global food production. Nat. Clim. Chang. 3, 904–908.
IPCC, 2014. Africa, in: Niang, I., Ruppel, O.C., Abdrabo, M.A., Essel, A., Lennard, C., Padgham, J., Urquhart, P. (Eds.), Impacts, Adaptation and Vulnerability - Contributions of the Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. pp. 1199–1265.
Kelley, C.P., Mohtadi, S., Cane, M.A., Seager, R., Kushnir, Y., 2015. Climate change in the Fertile Crescent and implications of the recent Syrian drought. Proc. Natl. Acad. Sci. 112, 3241–3246.
Knight, R.S., 2010. Statutory recognition of customary land rights in Africa - An investigation into best practices for lawmaking and implementation. FAO Legis. Study 105.
Knox, J., Hess, T., Daccache, A., Wheeler, T., 2012. Climate change impacts on crop productivity in Africa and South Asia. Environ. Res. Lett. 7, 34032.
Kusunose, Y., Mahmood, R., 2016. Imperfect forecasts and decision making in agriculture. Agric. Syst. 146, 103–110.
Leblois, A., Quirion, P., 2013. Agricultural insurances based on meteorological indices: realizations, methods and research challenges. Meteorol. Appl. 20, 1–9.
Lee, B.-H., Kenkel, P., Brorsen, B.W., 2013. Pre-harvest forecasting of county wheat yield and wheat quality using weather information. Agric. For. Meteorol. 168, 26–35.
Lesk, C., Rowhani, P., Ramankutty, N., 2016. Influence of extreme weather disasters on global crop production. Nature 529, 84–87.
Linnerooth-Bayer, J., Mechler, R., Hochrainer-Stigler, S., 2011. Insurance against Losses from Natural Disasters in Developing Countries. J. Integr. Disaster Risk Manag. 1, 59–81.
Lipper, L., Thornton, P., Campbell, B.M., Baedeker, T., Braimoh, A., Bwalya, M., Caron, P., Cattaneo, A.,

Garrity, D., Henry, K., Hottle, R., Jackson, L., Jarvis, A., Kossam, F., Mann, W., McCarthy, N., Meybeck, A., Neufeldt, H., Remington, T., Sen, P.T., Sessa, R., Shula, R., Tibu, A., Torquebiau, E.F., 2014. Climate-smart agriculture for food security. Nat. Clim. Chang. 4, 1068–1072.
Liu, B., Asseng, S., Müller, C., Ewert, F., Elliott, J., Lobell, D.B., Martre, P., Ruane, A.C., Wallach, D., Jones, J.W., Rosenzweig, C., Aggarwal, P.K., Alderman, P.D., Anothai, J., Basso, B., Biernath, C., Cammarano, D., Challinor, A., Deryng, D., Sanctis, G. De, Doltra, J., Fereres, E., Folberth, C., Garcia-Vila, M., Gayler, S., Hoogenboom, G., Hunt, L.A., Izaurralde, R.C., Jabloun, M., Jones, C.D., Kersebaum, K.C., Kimball, B.A., Koehler, A.-K., Kumar, S.N., Nendel, C., O'Leary, G.J., Olesen, J.E., Ottman, M.J., Palosuo, T., Prasad, P.V.V., Priesack, E., Pugh, T.A.M., Reynolds, M., Rezaei, E.E., Rötter, R.P., Schmid, E., Semenov, M.A., Shcherbak, I., Stehfest, E., Stöckle, C.O., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Thorburn, P., Waha, K., Wall, G.W., Wang, E., White, J.W., Wolf, J., Zhao, Z., Zhu, Y., 2016. Similar estimates of temperature impacts on global wheat yield by three independent methods. Nat. Clim. Chang. 6, 1130–1136.
Lobell, D.B., 2013. Errors in climate datasets and their effects on statistical crop models. Agric. For. Meteorol. 170, 58–66.
Lobell, D.B., Asseng, S., 2017. Comparing estimates of climate change impacts from process- based and statistical crop models. Environ. Res. Lett. 12, 1–12.
Lobell, D.B., Bänziger, M., Magorokosho, C., Vivek, B., 2011. Nonlinear heat effects on African maize as evidenced by historical yield trials. Nat. Clim. Chang. 1, 42–45.
Lobell, D.B., Burke, M.B., 2010. On the use of statistical models to predict crop yield responses to climate change. Agric. For. Meteorol. 150, 1443–1452.
Lobell, D.B., Hammer, G.L., McLean, G., Messina, C., Roberts, M.J., Schlenker, W., 2013. The critical role of extreme heat for maize production in the United States. Nat. Clim. Chang. 3, 497–501.
Lobell, D.B., Ortiz-Monasterio, J.I., Asner, G.P., Matson, P. a., Naylor, R.L., Falcon, W.P., 2005. Analysis of wheat yield and climatic trends in Mexico. F. Crop. Res. 94, 250–256.
Lobell, D.B., Roberts, M.J., Schlenker, W., Braun, N., Little, B.B., Rejesus, R.M., Hammer, G.L., 2014. Greater sensitivity to drought accompanies maize yield increase in the U.S. Midwest. Science 344, 516–9.
Lobell, D.B., Sibley, A., Ivan Ortiz-Monasterio, J., 2012. Extreme heat effects on wheat senescence in India. Nat. Clim. Chang. 2, 186–189.
Ma, S., Churkina, G., Trusilova, K., 2012. Investigating the impact of climate change on crop phenological events in Europe with a phenology model. Int. J. Biometeorol. 56, 749–763.
MAFSC, 2010. Agricultural statistics. Ministry of Agriculture, Food Security and Cooperatives . URL http://www.kilimo.go.tz/agricultural statistics/ (accessed 10.17.16).
MARS, 2017. Monitoring Agricultural ResourceS (MARS). URL https://ec.europa.eu/jrc/en/mars
McClung, C.R., 2014. Making hunger yield. Science 344, 699–700.
McIntosh, C., Sarris, A., Papadopoulos, F., 2013. Productivity, credit, risk, and the demand for weather index insurance in smallholder agriculture in Ethiopia. Agric. Econ. 44, 399–417.
Meze-Hausken, E., Patt, A., Fritz, S., 2009. Reducing climate risk for micro-insurance providers in Africa: A case study of Ethiopia. Glob. Environ. Chang. 19, 66–73.
Miao, R., Khanna, M., Huang, H., 2016. Responsiveness of Crop Yield and Acreage to Prices and Climate. Am. J. Agric. Econ. 98, 191–211.
Moore, F.C., Lobell, D.B., 2014. Adaptation potential of European agriculture in response to climate change. Nat. Clim. Chang. advance on.
Mueller, N.D., Gerber, J.S., Johnston, M., Ray, D.K., Ramankutty, N., Foley, J. a, 2012. Closing yield gaps through nutrient and water management. Nature 490, 254–257.
Müller, C., 2011. Agriculture: Harvesting from uncertainties. Nat. Clim. Chang. 1, 253–254.
Müller, C., Cramer, W., Hare, W.L., Lotze-Campen, H., 2011. Climate change risks for African agriculture. Proc. Natl. Acad. Sci. 108, 4313–4315.
Müller, C., Elliott, J., Chryssanthacopoulos, J., Arneth, A., Balkovic, J., Ciais, P., Deryng, D., Folberth, C., Glotter, M., Hoek, S., Iizumi, T., Izaurralde, R.C., Jones, C., Khabarov, N., Lawrence, P., Liu, W., Olin, S., Pugh, T.A.M., Ray, D., Reddy, A., Rosenzweig, C., Ruane, A.C., Sakurai, G., Schmid, E., Skalsky, R., Song, C.X., Wang, X., de Wit, A., Yang, H., 2016. Global Gridded Crop Model evaluation: benchmarking, skills, deficiencies and implications. Geosci. Model Dev. Discuss. 1–39.
Oury, B., 1965. Allowing for Weather in Crop Production Model Building. J. Farm Econ. 47, 270.
Phalkey, R.K., Aranda-Jan, C., Marx, S., Höfle, B., Sauerborn, R., 2015. Systematic review of current efforts to quantify the impacts of climate change on undernutrition. Proc. Natl. Acad. Sci. 201409769.
Qian, B., De Jong, R., Warren, R., Chipanshi, A., Hill, H., 2009. Statistical spring wheat yield forecasting for the Canadian prairie provinces. Agric. For. Meteorol. 149, 1022–1031.
Ramirez-Villegas, J., Challinor, A., 2012. Assessing relevant climate data for agricultural applications. Agric. For. Meteorol. 161, 26–45.
Ray, D.K., Gerber, J.S., MacDonald, G.K., West, P.C., 2015. Climate variation explains a third of global crop yield variability. Nat. Commun. 6, 1–9.

Ritchie, J.T., 1972. Model for predicting evaporation from row crop with incomplete cover. Water Resour. Res. 8, 1204–1213.
Rosenzweig, C., Elliott, J., Deryng, D., Ruane, A.C., Müller, C., Arneth, A., Boote, K.J., Folberth, C., Glotter, M., Khabarov, N., Neumann, K., Piontek, F., Pugh, T. a M., Schmid, E., Stehfest, E., Yang, H., Jones, J.W., 2014. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. Proc. Natl. Acad. Sci. 111, 3268–3273.
Rosenzweig, C., Iglesias, A., Yang, X.B., Epstein, P.R., Chivian, E., 2001. Climate change and extreme weather events: Implications for food production, plant diseases, and pests. Glob. Chang. Hum. Heal. 2, 90–104.
Rosenzweig, C., Tubiello, F.N., Goldberg, R., Mills, E., Bloomfield, J., 2002. Increased crop damage in the US from excess precipitation under climate change. Glob. Environ. Chang. 12, 197–202.
Rötter, R., Geijn, S. Van de, 1999. Climate change effects on plant growth, crop yield and livestock. Clim. Chang. 43, 651–681.
Rötter, R., Van de Geijn, S.C., 1999. Climate change effects on plant growth, crop yield and livestock. Clim. Chang. 43, 651–681.
Rötter, R.P., Carter, T.R., Olesen, J.E., Porter, J.R., 2011. Crop–climate models need an overhaul. Nat. Clim. Chang. 1, 175–177.
Rowhani, P., Lobell, D.B., Linderman, M., Ramankutty, N., 2011. Climate variability and crop production in Tanzania. Agric. For. Meteorol. 151, 449–460.
Sánchez, P.A., 2010. Tripling crop yields in tropical Africa. Nat. Geosci. 3, 299–300.
Sarris, A., 2013. Weather index insurance for agricultural development: introduction and overview. Agric. Econ. 44, 381–384.
Schauberger, B., Gornott, C., Wechsung, F., 2017. Global evaluation of a semiempirical model for yield anomalies and application to within-season yield forecasting. Glob. Chang. Biol.
Schewe, J., Otto, C., Frieler, K., 2017. The role of storage dynamics in annual wheat prices. Environ. Res. Lett. 12, 54005.
Schleussner, C.-F., Donges, J.F., Donner, R. V., Schellnhuber, H.J., 2016. Armed-conflict risks enhanced by climate-related disasters in ethnically fractionalized countries. Proc. Natl. Acad. Sci. 113, 9216–9221.
Semenov, M. a, Shewry, P.R., 2011. Modelling predicts that heat stress, not drought, will increase vulnerability of wheat in Europe. Sci. Rep. 1, 66.
Sileshi, G., Akinnifesi, F.K., Debusho, L.K., Beedy, T., Ajayi, O.C., Mong'omba, S., 2010. Variation in maize yield gaps with plant nutrient inputs, soil type and climate across sub-Saharan Africa. F. Crop. Res. 116, 1–13.
Snapp, S., Kerr, R.B., Smith, A., Ollenburger, M., 2014. Modeling and participatory farmer-led approaches to food security in a changing world: A case study from Malawi. Sécheresse 24, 350–358.
Statistical Offices of the Federation and the Länder, 2016. Datensatz Anbaufläche(Feldfrüchte und Grünland): Deutschland, Jahre, Fruchtarten . URL https://www.destatis.de/
Stoltzfus, J.D., Carter, J.Y., Akpinar-Elci, M., Matu, M., Kimotho, V., Giganti, M.J., Langat, D., Elci, O., 2014. Interaction between climatic, environmental, and demographic factors on cholera outbreaks in Kenya. Infect. Dis. Poverty 3, 37.
Stone, R.C., Meinke, H., 2005. Operational seasonal forecasting of crop performance. Philos. Trans. R. Soc. B Biol. Sci. 360, 2109–2124.
Surminski, S., Bouwer, L.M., Linnerooth-Bayer, J., 2016. How insurance can support climate resilience. Nat. Clim. Chang. 6, 333–334.
Thornton, P.K., Jones, P.G., Alagarswamy, G., Andresen, J., 2009. Spatial variation of crop yield response to climate change in East Africa. Glob. Environ. Chang. 19, 54–65.
Thornton, P.K., Jones, P.G., Alagarswamy, G., Andresen, J., Herrero, M., 2010. Adapting to climate change: Agricultural system and household impacts in East Africa. Agric. Syst. 103, 73–82.
Thornton, P.K., Jones, P.G., Ericksen, P.J., Challinor, A.J., 2011. Agriculture and food systems in Sub-Saharan Africa in a 4 C+ world. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 369, 117–136.
Tilman, D., Balzer, C., Hill, J., Befort, B.L., 2011. Global food demand and the sustainable intensification of agriculture. Proc. Natl. Acad. Sci. U. S. A. 108, 20260–4.
Tittonell, P., Giller, K.E., 2013. When yield gaps are poverty traps: The paradigm of ecological intensification in African smallholder agriculture. F. Crop. Res. 143, 76–90.
USDA, 2017. National Agricultural Statistics Service (NASS) . URL https://www.nass.usda.gov/Surveys/Guide_to_NASS_Surveys/Agricultural_Yield/
van der Velde, M., Folberth, C., Balkovič, J., Ciais, P., Fritz, S., Janssens, I. a., Obersteiner, M., See, L., Skalský, R., Xiong, W., Peñuelas, J., 2014. African crop yield reductions due to increasingly unbalanced nitrogen and phosphorus consumption. Glob. Chang. Biol. 20, 1278–1288.
van Dijk, M., Morley, T., Jongeneel, R., van Ittersum, M., Reidsma, P., Ruben, R., 2017. Disentangling agronomic and economic yield gaps: An integrated framework and application. Agric. Syst. 154, 90–99.
van Ittersum, M.K., Cassman, K.G., Grassini, P., Wolf, J., Tittonell, P., Hochman, Z., 2013. Yield gap analysis with local to global relevance—A review. F. Crop. Res. 143, 4–17.

Van Wart, J., Kersebaum, K.C., Peng, S., Milner, M., Cassman, K.G., 2013. Estimating crop yield potential at regional to national scales. F. Crop. Res. 143, 34–43.
Vitousek, P.M., Naylor, R., Crews, T., David, M.B., Drinkwater, L.E., Holland, E., Johnes, P.J., Katzenberger, J., Martinelli, L.A., Matson, P.A., Nziguheba, G., Ojima, D., Palm, C.A., Robertson, G.P., Sanchez, P.A., Townsend, A.R., Zhang, F.S., 2009. Nutrient Imbalances in Agricultural Development. Science (80). 324, 1519–1520.
Ward, P.S., Florax, R.J.G.M., Flores-Lagunes, A., 2014. Climate change and agricultural productivity in Sub-Saharan Africa: a spatial sample selection model. Eur. Rev. Agric. Econ. 41, 199–226.
Weedon, G.P., Balsamo, G., Bellouin, N., Gomes, S., Best, M.J., Viterbo, P., 2014. Data methodology applied to ERA-Interim reanalysis data. Water Resour. Res. 50, 7505–7514.
Welch, J.R., Vincent, J.R., Auffhammer, M., Moya, P.F., Dobermann, A., Dawe, D., 2010. Rice yields in tropical/subtropical Asia exhibit large but opposing sensitivities to minimum and maximum temperatures. Proc. Natl. Acad. Sci. 107, 14562–14567.
West, P.C., Gerber, J.S., Engstrom, P.M., Mueller, N.D., Brauman, K. a., Carlson, K.M., Cassidy, E.S., Johnston, M., MacDonald, G.K., Ray, D.K., Siebert, S., 2014. Leverage points for improving global food security and the environment. Science 345, 325–8.
Westengen, O.T., Ring, K.H., Berg, P.R., Brysting, A.K., 2014. Modern maize varieties going local in the semi-arid zone in Tanzania. BMC Evol. Biol. 14, 1.
Wheeler, T., von Braun, J., 2013. Climate Change Impacts on Global Food Security. Science (80). 341, 508–513.
Woodard, J.D., Garcia, P., 2008. Weather derivatives, spatial aggregation, and systemic risk: Implications for reinsurance hedging. J. Agric. Resour. Econ. 33, 34–51.
World Bank, 2017. World Bank Open Data - Agricultural land (% of land area), Arable land (% of land area), Tanzania . URL https://data.worldbank.org/
World Resource Institute, 2010. The precarious position of Tanzania's village land . URL https://agriknowledge.org/downloads/mg74qm138
Wouterse, F., 2010. Migration and technical efficiency in cereal production: evidence from Burkina Faso. Agric. Econ. 41, 385–395.
WTO, 2017. Domestic support . URL https://www.wto.org/english/tratop_e/agric_e/ag_intro03_domestic_e.htm
You, L., Ringler, C., Wood-Sichra, U., Robertson, R., Wood, S., Zhu, T., Nelson, G., Guo, Z., Sun, Y., 2011. What is the irrigation potential for Africa? A combined biophysical and socioeconomic approach. Food Policy 36, 770–782.
You, L., Rosegrant, M.W., Wood, S., Sun, D., 2009. Impact of growing season temperature on wheat productivity in China. Agric. For. Meteorol. 149, 1009–1014.

11 Supplemental Information

11.1 How do inputs and weather drive wheat yield volatility? The example of Germany

Hakon Albers[1*], Christoph Gornott[2], Silke Hüttel[3]

[1] Martin Luther University Halle-Wittenberg

[2] Potsdam Institute for Climate Impact Research (PIK)

[3] University of Rostock, Agricultural Economics

[*] Corresponding author

11.1.1 Literature Review

First, we discuss how weather can be included in production functions. Second, overview tables introduce the reviewed papers in a condensed way.

11.1.1.1 Three choices

Three choices are important for including weather into the production function framework: the selection of weather variables, aggregation levels of weather data and the functional form describing the input-output and weather-yield relationships.

Among the ***weather variables***, the majority employs temperature and precipitation, where radiation and evapotranspiration are also found, though sparsely. Variables capturing soil moisture are rarely applied since this requires spatially detailed data usually not available (e.g., Bakker et al., 2005). More recent papers emphasize vapor pressure deficit (VPD) as an important yield-determining variable (e.g., Lobell et al., 2014; Roberts et al., 2013). However, similar to other applied evapotranspiration variables, VPD does not include the water-holding capacity of soil and might thus indicate dry conditions in cases where water supply is sufficient. In addition, atmospheric CO_2 concentration has been proven to be important for yield levels, but in the literature we reviewed its small variation over time and space prevented researchers from quantifying its impacts in econometric models (e.g., Finger and Schmid, 2008, p. 26) with only few exceptions (e.g., Blanc, 2012).

As agronomic knowledge suggests, to ***aggregate data*** at different phenological stages improves evidence of weather effects on yield levels/ variability, although data requirements are high (Dixon et al., 1994). For instance, Butler and Huybers (2015) or Ortiz-Bobea and Just (2013) incorporate phenological stages while assessing impacts of weather conditions on maize yields to analyze adaptation possibilities of farms. While these studies emphasize temporal aggregation, spatial aggregation has been analyzed in greater detail by Garcia et al. (1987).

Finding the appropriate ***functional form*** describing the input-output relationship is an ongoing research topic, particularly in production economics literature (e.g., Coelli et al., 2005, Griffin et al., 1987). Indeed, an appropriate functional form is crucial since biases due to functional form misspecification are in the same dimension as omitted variable biases. Typically applied functional forms include Cobb-Douglas, linear, or quadratic with linear-additive weather effects. Other forms, particularly those accounting for non-linear weather effects, are rarely taken into account; exceptions are Odening et al. (2007), Schlenker and Roberts (2009) and Lobell et al. (2014).

11.1.1.2 Literature overview tables

In order to organize the vast literature on non-experimental cereal yield data (or output, excluding rice) in a convenient way, we produced overview tables which allow to compare several articles at once. A meta-table (S1) summarizes information from the overview tables (S3–S5). The order within the tables is chronological.

Tab. S1. Meta analysis of literature review

Characteristic	Absolute number	Share [%]
Total number of studies	37	100
Species maize	24	65
Species wheat	19	51
Variabe class evapotranspiration	8	22
Variabe class soil moisture	8	22
Aggregation over time: phenological	4	11
Total number of studies since 2010	16	100
Using production function approach	3	19
Using economic variables	5	31

Note: Source: own representation. Not all categories are exclusive. E.g., some studies analyse both maize and wheat, which explains why "maize" and "wheat" sum up to more than 100 percent.

Particularly, the row "PFA" (production function approach) addresses whether inputs are included. The criterion for being classified as "PFA: yes" was: The empirical model includes at least one variable measuring or approximating inputs, and can be classified as one category of the KLEMS approach (capital, labor, energy, material inputs and services; Coelli et al., 2005, p. 141) or as land. Furthermore, for each article, the tables provide characteristics such as time period, geographical focus, information on data aggregation and detrending, estimation technique, and included weather variables. In specific, several variables are used to quantify the process of evapotranspiration. These variables are collapsed into the group of variables called "e." Similarly, variables approximating the water-holding capacity of soil are grouped into the class "sm."

Tab. S2. Abbreviations literature overview tables

	meaning		meaning		meaning
agg	aggregation	im	implicit	qu	quadratic
AI	Ångström-Index	it	interaction terms	r	variables which measure
avg	average	irr	irrigated		or proxy radiation
cbnr	considered but not relevant	IV	instrumental variable	RCM	random coefficents
cc	cloud cover	lfa	less-favoured area		model
clr	crop land relevant	loc reg	local regression	RDI	Rainfall Deficit Index
cor	correlation	lsdv	least squares	re	random effects
Cobb-D.	Cobb-Douglas		dummy variable	reg	regional
CR	cumulative rainfall index	MARS	multivariate adaptive	rel	relative
cs	cross section		regression splines	RESET	regression specification
CV	coefficient of variation	max	maximum		error test
d	dummy variables	min	minimum	RSI	Rainfall Sum Index
DD	degree days	ml	maximum likelihood	Sept	September
DDD	damaging degree days	MPSMD	maximum potential	sh	sunshine hours
dev	deviation		soil moisture deficit	sm	variables which measure
DTR	diurnal temperature range	NAO	North Atlantic Oscillation		or proxy soil moisture
dyn	dynamic	NA	information not available	sp	spring
e(pot)	(potential) evapo-	nat	national	sqrt	squareroot
	transpiration	nls	non-linear least squares	sr	solar radiation
ex	explicit	Nov	November	SSA	Sub-Saharan Africa
fc	field capacity	NUTS	nomenclature of units	SWAP	soil water available
fd	first differencing		for territorial statistics *		to plants capacity
ff	functional form	ols	ordinary least squares	t	variables which measure
FGLS	feasible generalised	org	organic farming		temperature or related
	least squares	p	precipitation		variables (GDD, GSL)
GDD	growing degree days	PDSI	Palmer Drought	tfe	time fixed effects
gr	global radiation		Severity Index	transc	transcendental
gs	growing season	PFA	article follows a prod-	ts	time series
	(fixed months)		uction function approach	v	variable(s)
GSL	growing season length	pheno	phenological growth	VPD	vapor pressure deficit
GWR	geographically weighted		stages acknowledged	w	winter
	regression	pl	piecewise linear	ws	weather station
HDD	extreme heat degree days	pu	political units	wtd	within transformed
hm	hydrometeorological	prov	provinces		data
ife	individual fixed effects	psp	pre-season precipitation		

Note: Source: own representation. * European Commission (2007, pp. 4–6).

Tab. S3. Literature review – main facts

Authors	Region(s) / Period(s) / Species	Dependent v. / PFA; land / econ.; inst. v.
Gornott/Wechsung (2016)	Germany: counties / 1991–2010 / silage maize, wheat	fd log(yield) / yes / yes (nat); land
Miao et al. (2016)	US: 1,144 counties / 1977–2007 / maize	yield / no / yes; d
Butler/Huybers (2015)	U.S.: counties of 17 states / 1981–2012 / maize	yield / no / no; no
Ray et al. (2015)	World: 13,500 pu / 1979–2008 / maize, wheat	yield / no / no; no
Cai et al. (2014)	U.S.: 958 counties / 2002–6 / maize	(1) yield; (2) 5-year-yield / no / no; tfe
Lobell et al. (2014)	U.S.: farms of 9 states / 1995–2012 / maize	yield / no / no; no
Ward et al. (2014)	SSA: grid (2,653 cells) / 1997–2003 / cereals	yield / yes; no / yes; no
Brown (2013)	Scotland / 1963–2005 / barley, oats, wheat	yield / no / no; no
Ortiz-Bobea/Just (2013)	U.S.: counties of 8 states / 1985–2005 / maize	yield / no / no; no
Osborne/Wheeler (2013)	World: 14 countries / 1961–2009 / maize, wheat	yield % changes / no / no; no
Roberts et al. (2013)	Illinois / 1950–2010 / maize	yield / no / no; no
Blanc (2012)	SSA: 37 countries / 1961-2002 / maize, millet	fd log(yield) / yes / yes; d, land
Heimfarth et al. (2012)	Germany: 22 farms / 1997–2009 / w wheat	yield / no / no; no
Lobell et al. (2011)	World: countries / 1980–2008 / maize, wheat	log(yield) / no / no; no
Li et al. (2010)	China: (1) 27 prov; (2): grid / 1978–2000 / sp, w wheat	yield / no / no; no
Schlenker/Lobell (2010)	SSA: countries / 1961–2002 / maize, sorghum, millet	log(yield) / no / yes; no
Schlenker/Roberts (2009)	U.S.: counties of 30 states / 1950–2005 / maize	log(yield) / no / no; no
You et al. (2009)	China: 22 prov / 1979–2000 / soft wheat	yield / yes; no / yes; d
Lobell et al. (2008)	12 regions / 1961–2002 / maize, sorghum, wheat	fd of yield; fd of output / no / no; no
Deschênes/Greenstone (2007)	U.S.: counties / 1987, 92, 97, 2002 / maize	yield / no / yes; no
Hussain/Mudasser (2007)	Pakistan: 2 regions / 1975/6–99/2000 / w wheat	log(yield) / no / no; no
Lobell (2007)	10 countries / 1961–2002 / maize, wheat	fd log(yield); cbnr: yield / no / no; no
Odening et al. (2007)	Northeast Germany / 1993–2005 / w wheat	yield / no / no; no
Reidsma et al. (2007)	51,843 farms of EU-15 / 1990, 1995, 2000 / barley, maize, wheat	yield / yes / yes; lfa, land
Isik/Devadoss (2006)	Idaho: 4 crop districts / 1939–2001 / sp barley and wheat	yield / no / no; no
Schlenker/Roberts (2006)	U.S.: 1839 counties of 24 states / 1950–2004 / maize	log(yield) / no / no; no
Bakker et al. (2005)	several EU states: NUTS-2 or 3 / 1970s–2000 / soft wheat	avg-yield; trend yield / no / yes; no
Lobell et al. (2005)	Mexico: 2 regions / 1988–2002 / durum, w wheat	fd of yield / no / no; no
Chen et al. (2004)	U.S.: 48 states / 1973–97 / maize, sorghum, w wheat	yield; log(yield) / yes / yes; land
Lobell/Asner (2003)	U.S.: 618 counties / 1982–98 / maize	trend yield / no / no; no
Carter/Zhang (1998)	China: 249 counties (5 hm-regions) / 1980, 85, 87–90 / cereals	yield / yes; no / yes; no
Kaufmann/Snell (1997)	U.S.: 78 counties of 8 states / 1969, 74, 78, 82, 87 / maize	yield / yes / yes; loan rates, land
Nicholls (1997)	Australia / 1952–92 / wheat (unspecified)	yield / no / no; no
Dixon et al. (1994)	Illinois: 9 regions / 1953–1990 / maize	yield / yes; no / yes; no
Kaylen et al. (1992)	U.S.: 10 regions / 1949–88 / maize	yield / no / yes; no
Yang et al. (1992)	North Dakota / 1929–88 / wheat: sp, durum	yield / yes / yes; land
Hansen (1991)	U.S.: 3,057 firms of 10 states / 1988–9 / maize	yield / yes; no / yes; no

Note: Source: own representation. Abbreviations according to table S2.

Tab. S4. Literature review – econometric details

Authors	Data structure / Detrending / Estimator	FF / RESET / Validation
Gornott/Wechsung (2016)	ts, panel / fd / ols; ml; RCM	Cobb-D / yes / yes
Miao et al. (2016)	panel / im (qu) / IV; ife	linear, qu, square root / no / no
Butler/Huybers (2015)	panel / im (linear) / ols	linear / no / no
Ray et al. (2015)	ts / ex (linear – cubic) / ols	qu; it / no / yes
Cai et al. (2014)	(1) panel; (2) cs / im (linear) or tfe / loc reg (GWR)	linear / no / yes
Lobell et al. (2014)	panel; seperate cs / im; ex / MARS	pl / no / yes
Ward et al. (2014)	cs / not relevant / (Non-) spatial Heckit; ols	qu / no / no
Brown (2013)	ts / fd / cor	linear / no / no
Ortiz-Bobea/Just (2013)	panel / im (qu) / lsdv (ife) or ols on wtd	linear, qu / no / no
Osborne/Wheeler (2013)	ts / percentage changes, fd / ols	linear / no / no
Roberts et al. (2013)	ts / ex (linear) / ols	qu; it / no / yes
Blanc (2012)	panel / fd / lsdv (ife and tfe)	qu, log; it, / no / yes
Heimfarth et al. (2012)	separate ts / ex (linear) / ols (cor)	linear / no / no
Lobell et al. (2011)	panel / im (country, qu) / lsdv (ife)	qu / no / no
Li et al. (2010)	separate ts / ex (linear) and fd / ols (cor)	linear / no / no
Schlenker/Lobell (2010)	panel / im (qu) / lsdv (ife) or ols on wtd	linear, qu, pl / no / no
Schlenker/Roberts (2009)	panel, cs, ts / im (qu) / lsdv (ife); cbnr: (tfe)	step function, pl, polynomial, qu / no / yes
You et al. (2009)	panel / im (linear, cbnr: qu) / ols (reg d)	Cobb-D / yes / no
Lobell et al. (2008)	ts / fd all v / ols	linear / no / no
Deschênes/Greenstone (2007)	panel / im (tfe) / lsdv (ife); state by year tfe	qu; it / no / no
Hussain/Mudasser (2007)	panel / im (linear) / lsdv (ife)	transc [cbnr: Generalised Cobb-D, qu] / no / no
Lobell (2007)	ts / fd [cbnr: spline trend] / ols	linear / no / no
Odening et al. (2007)	ts / NA / ols	Leontief [cbnr: logarithmic, qu] / no / no
Reidsma et al. (2007)	cs, multilevel / not relevant / ml, RCM	linear, qu / no / no
Isik/Devadoss (2006)	panel / im (linear) / ml	qu / no / no
Schlenker/Roberts (2006)	panel / im (linear) / lsdv (ife) or ols on wtd	polynomial, qu / no / no
Bakker et al. (2005)	cs / not relevant / ols	linear / no / yes
Lobell et al. (2005)	separate ts / fd all v / ols	linear / no / no
Chen et al. (2004)	panel / im (linear); ex (stochastic) / ml, (re)	linear, Cobb-D / no / no
Lobell/Asner (2003)	cs / NA / ols; cor	linear / no / no
Carter/Zhang (1998)	5 panels / NA / FGLS (including reg d)	Cobb-D / no / no
Kaufmann/Snell (1997)	pooled cs / im (linear) / ols	logarithmic, linear, qu / no / no
Nicholls (1997)	ts / fd / ols	linear / no / no
Dixon et al. (1994)	panel / NA / FGLS	linear / yes / yes
Kaylen et al. (1992)	ts / stochastic / state space model	linear; weather v qu / no / no
Yang et al. (1992)	ts / im (cubic) / ols	linear, qu / no / no
Hansen (1991)	cs / not relevant / ml (tobit model)	qu, log; it [cbnr.: Cobb-D] / no / no

Note: Source: own representation. Abbreviations according to table S2.

Tab. S5. Literature review – details weather

Authors	Weather: p, t, r / e / sm / other	agg. time / space (finer scale than region applied)
Gornott/Wechsung (2016)	p, t, r: t normalized sr / e: epot / no / no	gs: vegetative, reproductive / no
Miao et al. (2016)	p, t: G(H)DD, dev / no / no / p: pre gs	gs (partly): monthly / grid
Butler/Huybers (2015)	p, r: cbnr: sh; t: max, min, G(K)DD / no / no / cbnr: freezing days	pheno.: 4 stages / no
Ray et al. (2015)	p, t / no / no / no	gs (monthly avg) / harvested area grid
Cai et al. (2014)	p, t / no / no / no	gs for U.S. (avg of monthly avg) / no
Lobell et al. (2014)	p, t: min, max / e: VPD / sm: cbnr: PDSI / no	dyn gs (sowing), 30-day-avg / grid
Ward et al. (2014)	p: avg, CV; t: avg, DTR / no / no / no	gs: monthly avg / no
Brown (2013)	p, t: DD, r: sh / e: MPSMD / sm: as e / NAO	monthly, annual, w, summer / clr grid
Ortiz-Bobea/Just (2013)	p, t: G(D)DD / no / no / no	pheno.: 3 stages / clr grid
Osborne/Wheeler (2013)	p, t / no / no / no	dyn gs (crop calendar) / clr grid
Roberts et al. (2013)	p, t: G(H)DD / e: VPD / no / no	dyn gs (t) / clr grid
Blanc (2012)	p: SPI, flood, t: t, drought / e: Hargreaves / no / CO_2	calendar year / clr grid
Heimfarth et al. (2012)	p: CR / no / no / no	gs (partly): daily avg / ws max 200 km from farms
Lobell et al. (2011)	p, t: avg, min, max / no / no / no	dyn gs (crop calendar), avg of monthly avg / clr grid
Li et al. (2010)	separately: p; t / no / no / no	gs for China / no
Schlenker/Lobell (2010)	p, t: avg, G(D)DD / no / no / no	gs: see Lobell et al. (2008); / clr grid
Schlenker/Roberts (2009)	p, t (includes GDD) / no / no / no	gs: one for U.S.: / clr grid
You et al. (2009)	p, t, r: cc / no / no / no	gs on provincal level monthly avg / grid
Lobell et al. (2008)	p, t / no / no / no	gs: avg of monthly avg / clr grid
Deschênes/Greenstone (2007)	p, t: GDD / no / sm: moisture capacity / v soil characteristics	dyn gs (t) / grid
Hussain/Mudasser (2007)	p, t: GSL / no / no / no	gs: (fixed months, t determined) / no
Lobell (2007)	p, t: avg, DTR / no / no / no	gs: avg of monthly avg / clr grid
Odening et al. (2007)	p: RDI, RSI / no / no / no	gs (partly): daily avg / no
Reidsma et al. (2007)	p, t / no / no / no	gs: avg of monthly avg / grid
Isik/Devadoss (2006)	p, t / no / no / no	gs (partly): monthly avg, annual / no
Schlenker/Roberts (2006)	p, t / no / no / no	dyn gs (t) / clr grid
Bakker et al. (2005)	p, t, r: gr / e: epot / sm: soil depth, SWAP / no	calendar year / sm: grid
Lobell et al. (2005)	t: min, max, r: sr / no / no / no	gs (partly): daily avg / no
Chen et al. (2004)	p, t / no / no / no	gs (partly) / no
Lobell/Asner (2003)	separately: trend of p (t), r: sr / no / no / no	gs (partly) for U.S. / no
Carter/Zhang (1998)	/ e: AI / no / dev from sample mean	monthly / no
Kaufmann/Snell (1997)	p: avg, max stage 6; t: min, daily max / no / no / length stage 1, 2	pheno.: 8 stages / no
Nicholls (1997)	p, t: min, max / no / no / cbnr: CO_2	calendar year / no
Dixon et al. (1994)	p, t, r: sr / no / sm, / no	pheno.: 4 stages / no
Kaylen et al. (1992)	p, t: monthly avg / no / sm: pre-gs p / no	gs: monthly and divisional avg / area and clr weighted
Yang et al. (1992)	t, / no / sm: avg of p and psp / no	gs (partly) / no
Hansen (1991)	p, t (all monthly, 30-year-avg) / e: it p t / sm: cbnr: fc / erosion	gs (partly) / 83 weather districts

Note: Source: own representation. Abbreviations according to table S2.

11.1.2 Data

We provide further details on the different FADN data (TF-8 vs. TF-14), yield data and data sources for the variable *share spring wheat*, details on input and weather data (used deflators, data aggregation). Furthermore, we include additional formal weather variable deficit explain how we merge different data sets and provide background information on the CAP reforms.

11.1.2.1 TF-8 versus TF-14

Alternatively, we considered to analyze farm type 13 "Specialist cereals, oilseed and protein crops (COP) (European Commission, 2010, p. 49)" (TF-14 grouping). One could argue that the latter type of farm would provide more precise data on inputs applied in wheat production in Germany. Higher aggregated data as from the TF-8 type 1 "Specialist field crops" are influenced to a higher degree by the use of inputs in production of other crops than winter wheat. However, we opted for the much higher representativeness of the TF-8 data. TF-8 is considerably closer to actual national output (see Fig. S1; Data from Statistisches Bundesamt, 2014a; other data sources as in paper. Represented output is calculated by multiplying average output of wheat SE110N with number of represented farms SYS02.). The importance of wheat within these groups does not differ much though. For 1995–2009, the mean shares of wheat in utilized land (without fallows and set-aside) are 0.3079 (TF8) and 0.3373 (TF14). A remaining limitation which applies to both groupings is that the data do not allow determining exactly the amount of inputs applied in wheat production.

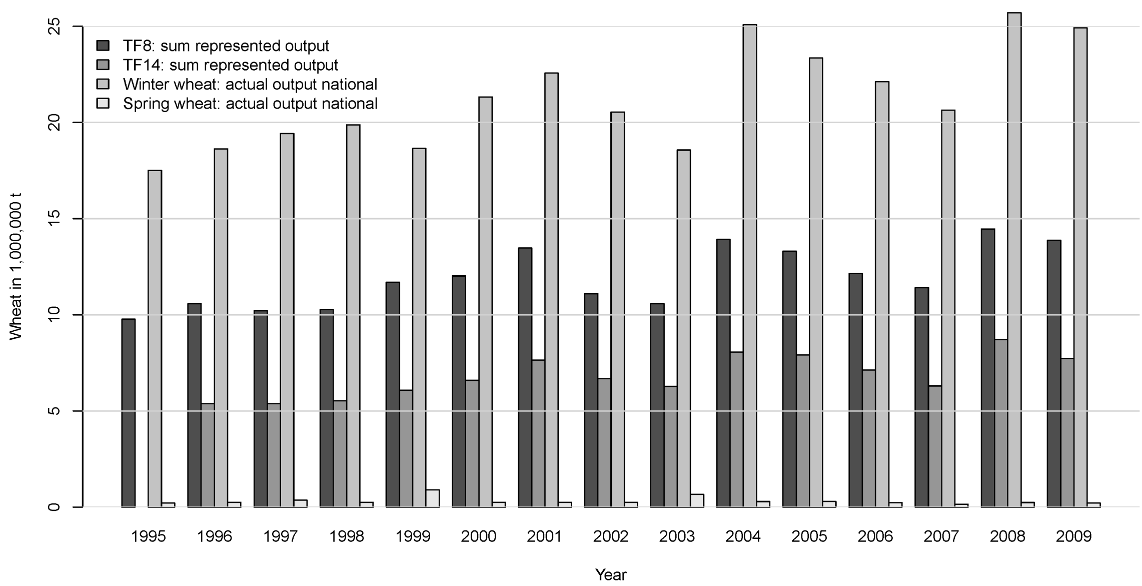


Fig. S1. Comparison of output represented by TF-8 and TF-14 grouping and actual national level output

11.1.2.2 Yields

We cannot rule out that wheat yields include production of spring wheat and spelt (European Commission, 2015). However, the contained share of spring wheat in total hectares of wheat can be regarded as low if the considered farms do not depart from the federal state trends. Spelt is of even lower acreage than spring wheat.

It can be argued that spring wheat might react differently to weather and application of inputs and that this might affect the results. For example, Chmielewski and Köhn (2000, p. 260) find different effects of weather on winter rye compared to spring cereals. Agricultural statistics show that the share of spring wheat in total wheat hectares (winter and spring wheat) was ca. 2% for 1995–2009, with a minimum of 0.3% and a maximum of 8.6%. We calculated these shares for each federal state and year using data from the German agricultural statistics (BMEL, 2015, various years). We included this variable as robustness check in the empirical model in order to account for regional differences and changes in spring wheat cultivation. The exact data sources are: BMEL (1995; 1996; 1997; 1998; 1999), BMVEL (2000; 2001; 2002), and BMELV (2005; 2008; 2009).

Data on acreage of spelt are neither available from German agricultural statistics nor from Statistisches Bundesamt nor from the federal state divisions of the latter, Statistische Ämter des Bundes und der Länder. A German non-scientific journal specializing in agriculture reports 22,833 hectares spelt harvested in Germany in 2004; most of the spelt is harvested in Bavaria and Baden-Wuerttemberg (Agrarzeitung online, 2004). For comparison: winter wheat amounted to 3,046,000 hectares and spring wheat to 48,300 hectares (sum of acreage for German federal states in 2004). We refrained from using specified winter wheat yield data which are available from German agricultural statistics instead of the FADN yields. This would result is a loss of information of which yields match which inputs and other farm characteristics. In addition, we would not have been able to determine the threshold of altitude as discussed below.

11.1.2.3 Details on input and weather data

We apply indices (base year 2010) without value-added tax (VAT) for the accounting year July until June (European Commission, 2010, p. 58; Statistisches Bundesamt, 2014b; 2014c, pp. 4, 16). VAT is not considered because all data are recorded without VAT (definition variable SE395, European Commission, 2015). Used price indices differ by input to be as precise as possible (Coelli et al., 2005, p. 155).

Capital: We approximate capital service flow, which would be the optimal measure according to Coelli et al. (2005, pp. 144–151) by aggregating two variables. (1) Depreciation based on replacement value (SE360). Used deflator equipment (contains machinery, tractors, buildings) and services for agricultural investment (series: *Waren und Dienstleistungen landwirtschaftlicher Investitionen*). (2) Machinery and building current costs (SE340); used deflator price index based on price indices for maintenance of both machinery and buildings (series: *Instandhaltung von Maschinen und Material ; Instandhaltung von Bauten*); applied weights (ca. 0.75 and 0.25, respectively) deduced from original weights in the complete index (Statistisches Bundesamt, 2014c, p. 10, Tab. 3).

Labor: total labor input in hours (SE011).
Land: for wheat in hectares (SE110D).

Energy: SE345, contains "motor fuels and lubricants, electricity, heating fuels (European Commission, 2015);" used deflator price index for energy and lubricants (series: *Energie und Schmierstoffe zusammen*).

Material inputs: aggregated from 2 variables. (1) Fertilizer (SE295), used deflator: price index all fertilizers (series: *Düngemittel zusammen*); (2) Crop protection (SE300), used deflator price index of all types of crop protection (series: *Pflanzenschutzmittel zusammen*).

Seeds: SE285, used deflator price index of seeds and plants (series: *Saatund Pflanzgut*).
Services: approximated with contract work (SE350): "Costs linked to work carried out by contractors and to the hire of machinery (European Commission, 2015)." There is no ready to use deflator we applied equally weighted price indices of maintenance of machinery and fuel (series: *Instandhaltung von Maschinen und Material, Treibstoffe zusammen*).
Manure: approximated by total livestock units (SE080).

Potential evapotranspiration: according to Turc-Ivanov (ETP_{TI}) calculated following Conradt et al. (2013, pp. 2950–2951, equation 1). ETP_{TI} considers average temperature, solar radiation, relative humidity (in %) and treats temperatures below and above 5°C differently. This measure has been superior over the evapotranspiration measure according to Haude (Haude, 1955).

Flood 2002: The Oder river flood affected several federal states: Bavaria, Brandenburg, Lower Saxony, Mecklenburg-Western Pomerania, Saxony, Saxony-Anhalt, Schleswig-Holstein, and Thuringia.

Other crop specific costs (SE305) are not considered, because they are a mixture of different costs which seem hardly relevant for this analysis. These include "[…] soil analysis, purchase of standing crops, renting crop land for a period of less than one year, purchase of crop products (grapes, etc.), costs incurred in the market preparation, storage, marketing of crops, etc. (European Commission, 2015)."

11.1.2.4 Weather: merging data sets and additional variable definitions

The daily weather variables (e.g., precipitation) are available for each weather station separately. Complex variables that are not observed, such as potential evapotranspiration, are calculated for each weather station using the observed weather variables (see following formulas). We then aggregate these daily weather variables (station level) to daily federal state averages.

The phenological data (not the same stations) are aggregated to federal state averages, that is, phenological stages differ by federal state and year. We then determine the days that are part of one of the four phenological periods as described in Tab. 1 in the paper. Finally, the daily weather variables (federal state level) are summed up for each of the phenological stages (federal state level).

Additional variable definitions

Days without precipitation (DWP):

$$DWP_p = \sum_{d=1}^{D} dwp_d = \begin{cases} 1, & \text{if } PREC_d = 0 \\ 0, & \text{if } PREC_d > 0 \end{cases}$$

with $PREC_d$ denoting the daily precipitation level and subscript d denoting a day within a phenological period p as deficit in the paper (Tab. 1).

Growing degree days (GDD):

$$GDD_p = \sum_{d=1}^{D} gdd_d = \begin{cases} Temp_{opt} - Temp_{min} & \text{if} \quad Temp_{avg,d} > Temp_{opt} \\ T_{avg,d} - T_{min} & \text{if} \quad Temp_{min} < Temp_{avg,d} \leq Temp_{opt} \\ 0 & \text{if} \quad Temp_{avg,d} \leq Temp_{min} \end{cases}$$

with $Temp_{opt}$ = 20◦C and $Temp_{min}$ = 4◦C. All temperatures refer to the daily average temperature ($Temp_{avg,d}$).

Killing degree days (KDD):

$$KDD_p = \sum_{d=1}^{D} kdd_d = \begin{cases} Temp_{avg,d} - Temp_{opt} & \text{if} \quad Temp_{avg,d} > Temp_{opt} \\ 0 & \text{if} \quad Temp_{avg,d} \leq Temp_{opt} \end{cases}$$

Potential evapotranspiration according to Haude (ETP_H) was calculated by the product of vapor pressure deficit (VPD) and an empirical correction factor, the Haude factor f_H (Haude, 1955). For f_H, we applied the factors for wheat used by Schrödter (1985). VPDd was calculated using the Magnus formula (Sonntag, 1990) with the maximum ($Temp_{max,d}$) and the minimum temperature ($Temp_{min,d}$) instead of the dew point temperature (Castellvi et al., 1996; 1997).

$$ETP_{H_p} = \sum_{d=1}^{D} f_{H_d} \underbrace{6.11\left(\exp\left(\frac{17.269\ Temp_{max,d}}{237.3 + Temp_{max,d}}\right) - \exp\left(\frac{17.269\ Temp_{min,d}}{237.3 + Temp_{min,d}}\right)\right)}_{=VPD_d}$$

Temperature normalized solar radiation: Similar as Gornott and Wechsung (2016, p. 92):

$$SRT_p = \sum_{d=1}^{D} \frac{SR_d}{Temp_{avg,d} + 20}$$

11.1.2.5 Altitude of farms, weather stations and phenological observational units

Merging different data sets must acknowledge the altitudes of farms, weather stations, and phenological observational units, because climate depends on altitude. Unfortunately, the publicly available FADN data do not provide information on altitudes at which the farms operate in Germany. Matching such information with data on the altitude of weather stations would be superior. Because this is not feasible, we do not consider weather stations above 600 m above sea level. The reason is that weather stations above this threshold are not relevant for the type of farm for which we analyze FADN data. According to IEEP (2006, p. 9), classification as Less Favoured Area (LFA) (Mountain Areas) in Germany applies to altitudes above 600m (with additional slope condition; see also European Commission, 2008, p. 4). According to an analysis of FADN data for the years 2004/05, farms specializing in field crops (TF-8 group 1, European Commission, 2010, p. 52) are not represented in areas classified as LFA Mountain Areas in Germany (European Commission, 2008, p. 61). Likewise we do not consider phenological observational units above 600m.

11.1.2.6 Details on CAP reforms

First, the Agenda 2000 reform, which was implemented in the same year, reduced product-specific price support and introduced area-based compensation payments. The payment level still depended on the planted crop. Biased incentives through still-coupled per-hectare payments may have led farmers to grow wheat as a major grain on marginal land. The midterm review of the Agenda 2000 in 2003 (Fischler-reform; active in 2005) led to payments being fully de-coupled from production in Germany. This may have also fostered incentives to plant wheat on marginal land. The Health Check of the CAP reforms in 2008 led to advanced reduction of the price support, and compulsory set-aside was abolished.

11.1.3 Regression model

The complete analysis was carried out using the software R (R Development Core Team, 2016). In the following, we provide information on model building and selection, robustness checks, and specification tests. In addition, we present additional plots and the cross-validation investigating the robustness of the regression model.

11.1.3.1 Model building, selection, specification tests and robustness

The following Tab. S6 shows the tested functional forms for the right-hand side of the model. The model with the right-hand side of the production function in quadratic form (that is, the inputs are in levels but the dependent variable is in logs) and weather in logs achieves the lowest Akaike Information Criterion (AIC).

Tab. S6 Results specification simplified models ordered by AIC

Production function	Specification weather	p-value RESET	AIC	R^2	k
Quadratic dv log	log	0.420	-393.970	0.835	49
Quadratic dv log	level	0.732	-389.550	0.839	53
Translog	level	0.300	-388.340	0.843	56
Translog	log	0.231	-371.600	0.846	66

Note: dv: dependent variable; k: number of parameters; RESET: regression specification error test.

Tab. S7 shows results of specification tests of the two models in Tab. 3 of the paper.

Tab. S7 Specification Tests

Test	p-values model 1	p-values model 2
Breusch-Pagan	0.301	0.769
Wooldridge's FD test h0=FE	0.002	0.896
Wooldridge's FD test h0=FD	$<$0.001	$<$0.001
RESET	0.49	0.834

Note: Croissant and Millo (2008, 28–9) provide details on Wooldridge's FD test.

The following three tables present several robustness checks. Tab. S8 shows which statistically insignificant variables were dropped after model building and selection (see step 4b, Fig. 2 in paper).

Tab. S8. Robustness check: Dropping statistically insignificant weather variables

	(1) Final specification	(3) from stepwise model search	(4) Drop stat. insign. weather var.
Intercept	0.0719 (0.0116)***	0.0883 (0.0078)***	0.0746 (0.0123)***
Inputs			
Capital	0.0433 (0.0399)	0.0413 (0.0406)	0.0451 (0.0406)
Labor	−0.1953 (0.1743)	−0.1881 (0.1902)	−0.1906 (0.1772)
Energy	0.1227 (0.0765)	0.1462 (0.0832)*	0.1247 (0.0769)
Material inputs	0.0911 (0.0305)***	0.0957 (0.0313)***	0.0879 (0.0308)***
Seeds	−0.1419 (0.0495)***	−0.1396 (0.0342)***	−0.1421 (0.0484)***
Manure	−5.5401 (6.9399)	−10.0438 (8.3737)	−6.5750 (6.8994)
Land wheat	−0.0495 (0.0418)	−0.0403 (0.0389)	−0.0475 (0.0434)
Capital squared	0.0007 (0.0002)***	0.0008 (0.0002)***	0.0007 (0.0002)***
Labor squared	0.0165 (0.0095)*	0.0179 (0.0108)*	0.0168 (0.0095)*
Energy squared	−0.0114 (0.0020)***	−0.0112 (0.0024)***	−0.0118 (0.0020)***
Material inputs squared	−0.0011 (0.0002)***	−0.0013 (0.0002)***	−0.0012 (0.0002)***
Seeds squared	0.0041 (0.0011)***	0.0040 (0.0008)***	0.0038 (0.0010)***
Capital * labor	−0.0044 (0.0015)***	−0.0045 (0.0015)***	−0.0044 (0.0015)***
Capital * seeds	−0.0008 (0.0001)***	−0.0009 (0.0002)***	−0.0008 (0.0001)***
Labor * energy	0.0111 (0.0036)***	0.0105 (0.0038)***	0.0113 (0.0037)***
Seeds * manure	0.2884 (0.1611)*	0.3210 (0.1607)**	0.2810 (0.1669)*
Weather			
Prec. stage 1	0.0499 (0.0159)***	0.0733 (0.0227)***	0.0545 (0.0163)***
Prec. stage 1 * East	−0.1675 (0.0439)***	−0.1542 (0.0477)***	−0.1600 (0.0437)***
Prec. stage 1 squared * East	−0.9129 (0.1412)***	−0.7915 (0.1949)***	−0.8648 (0.1333)***
Pot. evapotranspiration stage 1	0.2251 (0.0512)***	0.3017 (0.0430)***	0.2413 (0.0551)***
Growing degree days stage 2	0.1136 (0.0313)***	0.1881 (0.0425)***	0.1571 (0.0428)***
Solar radiation stage 2	−0.0232 (0.0423)	−0.0030 (0.0350)	−0.0260 (0.0420)
Solar radiation stage 2 squared	0.4425 (0.1561)***	0.6280 (0.1943)***	0.4576 (0.1495)***
Prec. stage 2	−0.0190 (0.0067)***	−0.0314 (0.0177)*	−0.0350 (0.0147)**
Killing degree days stage 3	−0.0182 (0.0027)***	−0.0171 (0.0030)***	−0.0183 (0.0028)***
Prec. stage 4	−0.0394 (0.0164)**	−0.0403 (0.0107)***	−0.0319 (0.0137)**
Prec. stage 4 * East	0.0843 (0.0519)	0.0903 (0.0484)*	0.0840 (0.0510)
Prec. stage 4 squared * East	0.2616 (0.1615)	0.2941 (0.1409)**	0.2946 (0.1587)*
Additional weather variables			
Prec. stage 2 squared * East		0.1682 (0.1116)	
Max. Temp. stage 2		−0.1540 (0.0876)*	−0.1247 (0.0941)
Temp. normalized radiation stage 4		0.1122 (0.1363)	
Pot. evapotranspiration stage 4		−0.0388 (0.0965)	
R^2	0.8290	0.8355	0.8303
Adj. R^2	0.7679	0.7691	0.7677

Note: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Dependent variable: first differences of logged wheat yield. Model 1 as in table 3 in paper. Controls and year dummy variables included but not shown. Weather in logs, inputs in levels. Coefficients/standard errors for inputs multiplied by 100. Explanatory variables first differenced; weather/inputs mean centered. SCC standard errors in () (Driscoll and Kraay 1998). Pot.: potential; prec.: precipitation; temp.: temperature.

Tab. S9. Robustness Check: Logged Inputs

	(1) Final specification	(5) log inputs
Intercept	0.0719 (0.0116)***	0.0744 (0.0123)***
Inputs		
Capital	0.0433 (0.0399)	0.3298 (0.1280)**
Labor	−0.1953 (0.1743)	−0.0797 (0.0631)
Energy	0.1227 (0.0765)	0.1216 (0.0744)
Material inputs	0.0911 (0.0305)***	0.1476 (0.0788)*
Seeds	−0.1419 (0.0495)***	−0.1275 (0.0430)***
Manure	−5.5401 (6.9399)	−0.0245 (0.0268)
Land wheat	−0.0495 (0.0418)	−0.0105 (0.0538)
Capital squared	0.0007 (0.0002)***	1.7877 (0.7489)**
Labor squared	0.0165 (0.0095)*	0.0359 (0.4125)
Energy squared	−0.0114 (0.0020)***	−3.9231 (0.7876)***
Material inputs squared	−0.0011 (0.0002)***	−0.7344 (0.1895)***
Seeds squared	0.0041 (0.0011)***	0.3143 (0.1695)*
Capital * labor	−0.0044 (0.0015)***	−0.9657 (0.5539)*
Capital * seeds	−0.0008 (0.0001)***	−0.3515 (0.0972)***
Labor * energy	0.0111 (0.0036)***	1.4662 (0.4706)***
Seeds * manure	0.2884 (0.1611)*	0.0963 (0.0845)
Weather		
Prec. stage 1	0.0499 (0.0159)***	0.0416 (0.0141)***
Prec. stage 1 * East	−0.1675 (0.0439)***	−0.1620 (0.0414)***
Prec. stage 1 squared * East	−0.9129 (0.1412)***	−0.7536 (0.1300)***
Pot. evapotranspiration stage 1	0.2251 (0.0512)***	0.2105 (0.0580)***
Growing degree days stage 2	0.1136 (0.0313)***	0.0923 (0.0423)**
Solar radiation stage 2	−0.0232 (0.0423)	−0.0152 (0.0539)
Solar radiation stage 2 squared	0.4425 (0.1561)***	0.1543 (0.1307)
Prec. stage 2	−0.0190 (0.0067)***	−0.0253 (0.0070)***
Killing degree days stage 3	−0.0182 (0.0027)***	−0.0155 (0.0032)***
Prec. stage 4	−0.0394 (0.0164)**	−0.0415 (0.0175)**
Prec. stage 4 * East	0.0843 (0.0519)	0.0716 (0.0525)
Prec. stage 4 squared * East	0.2616 (0.1615)	0.2484 (0.1835)
R^2	0.8290	0.8138
Adj. R^2	0.7679	0.7472

Note: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Dependent variable: first differences of logged wheat yield. Model 1 as in table 3 in paper. Controls and year dummy variables included but not shown. Weather in logs, inputs in levels (model 1) or in logs (model 5). Coefficients/standard errors for model 1 multiplied by 100. Explanatory variables first differenced; weather/inputs mean centered. SCC standard errors in () (Driscoll and Kraay, 1998). Pot.: potential; prec.: precipitation.

Tab. S9 provides a version of the final model with logged inputs (model 5). Model 5 passes specification tests such as the regression specification error test (RESET). However, model fit criteria for model 5 are slightly weaker compared to model 1 (R^2 is 0.83 and AIC: –396 for model 1; R^2 is 0.81 and AIC: -381 for model 5).

Moreover, the Davidson-MacKinnon-J-test (Wooldridge, 2009, p. 305) rejects the null hypothesis that model 1 cannot add explanatory power to model 5 at the 1% level, but the reverse is not true (based on robust covariance matrix). That is, the J-test favors model 1 over model 5.

Finally, Tab. S10 shows instrumental variable estimates. For illustration purposes, we rely on a simpler model where the dependent variable is still the difference of the logged yield but only the difference material inputs enter on the right-hand side (and the dummy variable for the extreme observation in Brandenburg 2003). We use the second lagged difference of material inputs (linear and squared term) as instruments. These are exogenous to yields in year *t*.

First-stage regressions show a statistically significant correlation (based on robust standard errors) of the difference of material inputs (linear and squared terms) and the IVs (both R^2 are roughly 0.2, irrespective of whether the Brandenburg dummy variable is included). Hence, the used instruments could be treated as valid.

The results are shown in the column labeled model 6. While the results do not qualitatively change, that is, the signs of material inputs are preserved within the IV results, we find a modest quantitative difference between the marginal effects of material inputs in model 1 and the IV estimate. The marginal effect in the IV estimate is about 1 percentage point larger at a 30 EUR increase of material inputs (2.8% compared to 1.7% in model 1), but also has a smaller positive range (due to the more prominent quadratic term). This difference—in our judgment—does not query the results of model 1. While the linear term is not significant alone at conventional levels ($p = 0.15$; robust standard errors), both linear and squared term are jointly significant (F-test; $p < 0.001$; robust standard errors). For further robustness checks, we have also included capital, where the interpretability of the results suffered from considerable increases in the standard errors, which is a known problem of IV regression (Wooldridge, 2009, p. 511).

Tab. S10. Robustness Check: Instrumental Variable Estimates

	(1) Final specification	(6) IV: material inputs	(7) IV: material inputs + capital	(8) IV: logged material inputs
Intercept	0.0719 (0.0116)***	0.0094 (0.0209)	0.0861 (0.0945)	0.0074 (0.0195)
Capital	0.0433 (0.0399)		0.4865 (0.5843)	
Material inputs	0.0911 (0.0305)***	0.2192 (0.1519)	0.2265 (0.1829)	0.3630 (0.3123)
Material inputs squared	−0.0011 (0.0002)***	−0.0042 (0.0019)**	−0.0046 (0.0022)**	−3.7625 (0.8254)***
Year 2003 * Brandenburg	−0.3877 (0.0229)***	−0.5639 (0.0340)***	−0.5679 (0.0461)***	−0.6301 (0.0759)***
R^2	0.8290	0.2593	0.1409	0.2649

Note: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Dependent variable: first differences of logged wheat yield. Model 1 as in table 3 in paper but other parameters not shown. Inputs in levels (except model 8). Coefficients/standard errors for models with inputs in levels multiplied by 100. Explanatory variables first differenced; weather/inputs mean centered. Used instruments: own second lagged difference. SCC (Driscoll and Kraay, 1998) standard errors in ().

11.1.3.2 Actual and fitted values

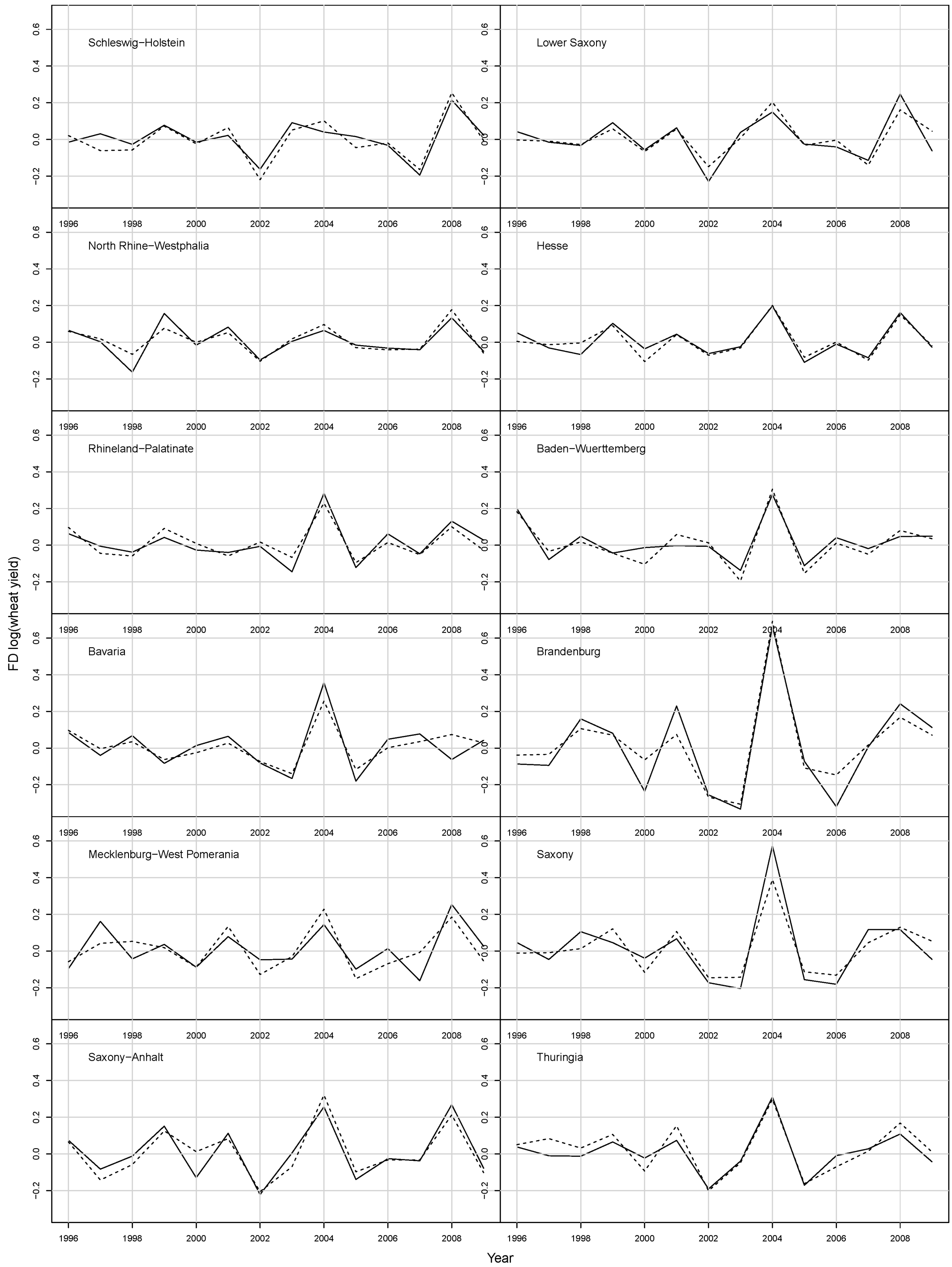


Fig. S2. Actual (solid line) and fitted (dashed) yield changes based on model 1

11.1.3.3 Cross validation

The average root mean squared error for the predicted years is 0.11. The Pearson correlation between actual and the out-of-sample predicted values is 0.51 ($R^2 = 0.26$). Notably, the out-of-sample prediction errors (RMSE, mean absolute error) are higher for yield changes following the heatwave year. That is, the model's predictive power is considerably weaker for 2004 and the Federal State of Brandenburg, which considerably suffered from the heatwave (Fig. S3). Within the prediction, we do not take time shocks into account and thus argue that the interaction of weather, inputs, and yield changes seems to be well-captured by the production function model.

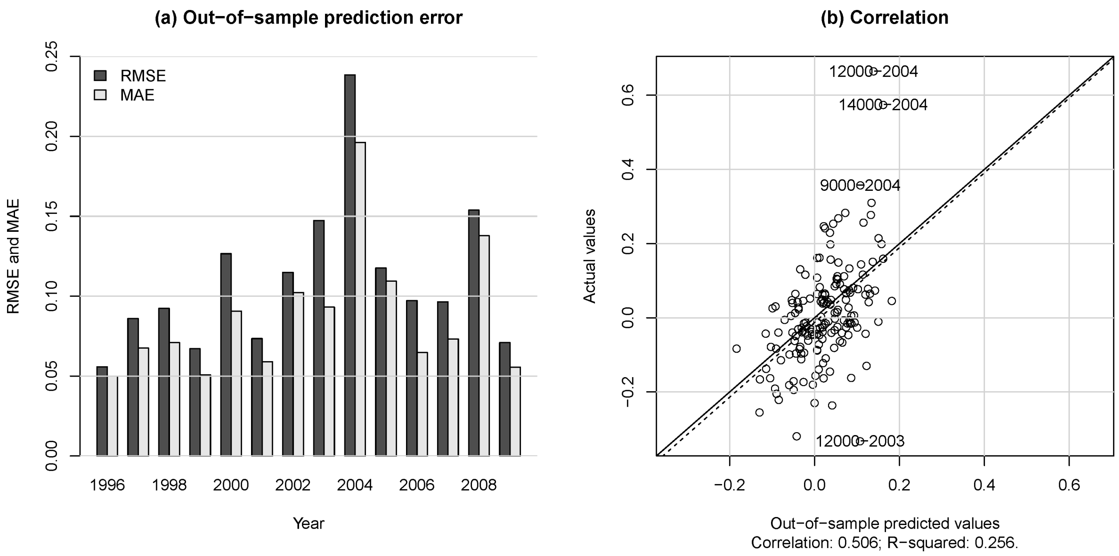


Fig. S3. (a) Prediction errors; (b) Actual and out-of-sample predicted values. Solid: perfect fit line; dashed: from linear regression

11.1.4 Yield volatility

The section contains the result table underlying the maps depicting the volatilities in the main paper and information on the robustness check regarding a possible aggregation bias.

11.1.4.1 Detailed results

Tab. S11. Actual, Weather- and Input-induced Yield Volatility [in %]: (1) 1996–2002 and (2) 2003–2009

					Volatility model 1				Volatility model 2	
	actual		weather		inputs		inputs + weather		weather	
Period	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Western Germany										
Schleswig-Holstein (SH)	7.52	12.38	3.23	3.72	7.64	5.12	6.61	7.90	2.88	4.26
Lower Saxony (LS)	10.70	12.84	5.50	4.04	2.40	3.26	5.88	5.21	5.15	5.05
North Rhine-Westphalia (NRW)	10.94	6.66	4.62	5.38	7.57	3.34	6.39	7.65	4.23	6.49
Hesse (HE)	6.55	11.85	5.57	4.63	2.14	3.50	6.11	4.20	4.41	4.53
Rhineland-Palatinate (RP)	3.97	14.99	5.46	4.11	5.17	5.05	7.53	3.61	4.93	4.87
Baden-Wuerttemberg (BW)	8.99	13.66	4.08	3.90	7.10	3.49	7.52	5.56	3.69	3.74
Bavaria (BV)	7.20	18.24	4.42	2.93	6.64	3.80	3.77	5.51	3.75	2.45
Eastern Germany										
Brandenburg (BB)	18.97	34.58	5.74	7.08	6.31	4.81	9.19	8.56	7.34	8.23
Mecklenburg-West Pomerania (MWP)	9.54	14.23	3.57	2.73	6.22	4.00	6.50	4.93	3.09	2.59
Saxony (SY)	9.43	27.42	3.83	6.55	5.16	7.10	7.30	10.95	4.34	7.07
Saxony-Anhalt (SA)	13.67	16.17	5.57	6.82	5.80	8.48	10.34	7.38	6.59	7.47
Thuringia (TH)	8.92	15.09	4.58	2.75	5.90	5.41	9.71	5.65	4.15	3.72

National level volatility based on yearly dummy variables: *model 1*: (1) 4.99, (2) 11.17; *model 2*: (1) 6.1, (2) 13.03.

11.1.4.2 Aggregation bias

Admittedly, spatial dependence in weather-related yield variability might not adhere to the administrative borders of federal states. In addition, the level of geographical aggregation differ by state; however, to acknowledge input changes as determinants of yield changes, a common level of aggregation for which all data are available must be applied. While we cannot circumvent different geographical aggregation, we must acknowledge that this might drive the results towards more heterogeneous volatilities across regions because in smaller geographical units the share of idiosyncratic factors in the variance might be higher due to the lower aggregation level. To investigate a possible aggregation bias, we run an auxiliary regression of weather-volatilities on hectares of wheat at the state level in each period. In the presence of aggregation bias, volatilities would decrease in the number of hectares planted with wheat. The estimated coefficient remains insignificant and we thus rule out any aggregation bias.

11.1.5 References

Agrarzeitung online, 10 2004. Dinkel profitiert von stetig wachsender Nachfrage. Accessed 12.04.2014, available at http: //www.agrarzeitung.de/nachrichten/wirtschaft/protected/ dinkel-profitiert-von-stetig-wachsender-nachfrage-15955.html.

Bakker, M. M., Govers, G., Ewert, F., Rounsevell, M., Jones, R., 2005. Variability in regional wheat yields as a function of climate, soil and economic variables: Assessing the risk of confounding. Agric. Ecosyst. Environ. 110 (3-4), 195–209.

Blanc, E., 2012. The impact of climate change on crop yields in Sub-Saharan Africa. Amer. J. Clim. Chang. 1 (1), 1–13.

BMEL, 1995. Getreideanbauflachen nach Getreidearten und Ländern, 1995. In: Statistik des Bundesministerium für Ernährung und Landwirtschaft. Bundesministerium für Ernährung und Landwirtschaft, accessed 10.02.2014, available at http://berichte.bmelv-statistik.de/EQT-0101010-1995.pdf.

BMEL, 1996. Getreideanbauflachen nach Getreidearten und Ländern, 1996. In: Statistik des Bundesministerium für Ernährung und Landwirtschaft. Bundesministerium für Ernährung und Landwirtschaft, accessed 10.02.2014, available at http://berichte.bmelv-statistik.de/EQT-0101010-1996.pdf.

BMEL, 1997. Getreideanbauflachen nach Getreidearten und Ländern, 1997. In: Statistik des Bundesministerium für Ernährung und Landwirtschaft. Bundesministerium für Ernährung und Landwirtschaft, accessed 10.02.2014, available at http://berichte.bmelv-statistik.de/EQT-0101010-1997.pdf.

BMEL, 1998. Getreideanbauflachen nach Getreidearten und Ländern, 1998. In: Statistik des Bundesministerium für Ernährung und Landwirtschaft. Bundesministerium für Ernährung und Landwirtschaft, accessed 10.02.2014, available at http://berichte.bmelv-statistik.de/EQT-0101010-1998.pdf.

BMEL, 1999. Getreideanbauflachen nach Getreidearten und Ländern, 1999. In: Statistik des Bundesministerium für Ernährung und Landwirtschaft. Bundesministerium für Ernährung und Landwirtschaft, accessed 10.02.2014, available at http://berichte.bmelv-statistik.de/EQT-0101010-1999.pdf.

BMEL, 2015. Statistik und Berichte des Bundesministerium für Ernährung und Land- wirtschaft. Accessed 23.03.2016, available at http://www.bmel-statistik.de/.

BMELV, 2005. Besondere Ernte- und Qualitätsermittlung (BEE) 2005. Accessed 09.02.2013, available at http://berichte.bmelv-statistik.de/ EQB-1002000-2005.pdf.

BMELV, 2008. Besondere Ernte- und Qualitätsermittlung (BEE) 2008. Accessed 10.02.2013, available at http://berichte.bmelv-statistik.de/ EQB-1002000-2008.pdf.

BMELV, 2009. Besondere Ernte- und Qualitätsermittlung (BEE) 2009. Accessed 10.02.2014, available at http://berichte.bmelv-statistik.de/ EQB-1002000-2009.pdf.

BMVEL, 2000. Abschlussbericht über die Besondere Ernteermittlung bei Getreide und Kartoffeln in Deutschland. Accessed 10.02.2014, available at http: //berichte.bmelv-statistik.de/EQB-1002000-2000.pdf.

BMVEL, 2001. Abschlussbericht über die Besondere Ernteermittlung bei Getreide und Kartoffeln in Deutschland. Accessed 09.02.2013, available at http: //berichte.bmelv-statistik.de/EQB-1002000-2001.pdf.

BMVEL, 2002. Abschlussbericht über die Besondere Ernteermittlung bei Getreide und Kartoffeln in Deutschland 2002. Erträge und Qualität. Accessed 10.02.2014, available at http://berichte.bmelv-statistik.de/EQB-1002000-2002.pdf.

Brown, I., 2013. Influence of seasonal weather and climate variability on crop yields in Scotland. Int. J. Biometeorol. 57 (4), 605–614.

Butler, E. E., Huybers, P., 2015. Variations in the sensitivity of US maize yield to extreme temperatures by region and growth phase. Environ. Res. Lett. 10 (3), 1–8.

Cai, R., Yu, D., Oppenheimer, M., 2014. Estimating the spatially varying responses of corn yields to weather variations using geographically weighted panel regression. J. Agr. Resource Econ. 39 (2), 230–252.

Carter, C. A., Zhang, B., 1998. The weather factor and variability in China's grain supply. J. Compar. Econ. 26 (3), 529–543.

Castellvi, F., Perez, P. J., Villar, J. M., Rose, J. I., 1996. Analysis of methods for estimating vapor pressure deficit and relative humidity. Agric. For. Meteorol. 82, 29–45.

Castellvi, F., Perez, P. J., Villar, J. M., Rose, J. I., 1997. Methods for estimating vapor pressure deficit at a regional scale depending on data availability. Agric. For. Meteorol. 87, 243–252.

Chen, C.-C., McCarl, B. A., Schimmelpfennig, D. E., 2004. Yield variability as influenced by climate: A statistical investigation. Clim. Chang. 66 (1-2), 239–261.

Chmielewski, F.-M., Köhn, W., 2000. Impact of weather on yield components of winter rye over 30 years. Agric. For. Meteorol. 102, 253–261.

Coelli, T. J., Rao, D. S. P., O'Donnell, C. J., Battese, G. E., 2005. An Introduction to Efficiency and Productivity Analysis, 2nd Edition. Springer, New York.

Conradt, T., Wechsung, F., Bronstert, A., 2013. Three perceptions of the evapotranspiration landscape: comparing spatial patterns from a distributed hydrological model, remotely sensed surface temperatures, and subbasin water balances. Hydrol. Earth Syst. Sci. 17 (7), 2947–2966.

Croissant, Y., Millo, G., 2008. Panel data econometrics in r: The plm package. J. Stat. Softw. 27 (2), 1–43.

Deschˆenes, O., Greenstone, M., 2007. The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather. Amer. Econ. Rev. 94 (1), 354–385.
Dixon, B. L., Hollinger, S. E., Garcia, P., Tirupattur, V., 1994. Estimating corn yield response models to predict impacts of climate change. J. Agr. Resource Econ. 19 (1), 58–68.
Driscoll, J. C., Kraay, A. C., 1998. Consistent covariance matrix estimation with spatially dependent panel data. Rev. Econ. Statist. 80, 549–560.
European Commission, 2007. European Regional and Urban Statistics. Reference Guide. Luxembourg.
European Commission, 2008. Overview of the less favoured areas farms in the EU- 25 (2004-2005). European Commission, Brussels, accessed 03.02.2014, available at http://ec.europa.eu/agriculture/rica/pdf/rd0101_overview_lfa.pdf.
European Commission, 2010. Farm Accounting Data Network. An A to Z of Methodology. Accessed 10.02.2014, available at http://ec.europa.eu/agriculture/ rica//pdf/site_en.pdf.
European Commission, 2015. FADN – Variable Definitions. Accessed 30.11.2015, available at http://ec.europa.eu/agriculture/rica//database/help/ infometa.csv.
Finger, R., Schmid, S., 2008. Modeling agricultural production risk and the adaption to climate change. Agr. Finance Rev. 68 (1), 25–41.
Garcia, P., Offutt, S. E., Pinar, M., September 1987. Corn yield behavior: Effects of technological advance and weather conditions. J. Climate Appl. Meteor. 26, 1092– 1102.
Gornott, C., Wechsung, F., 2016. Statistical regression models for assessing climate impacts on crop yields: A validation study for winter wheat and silage maize in Germany. Agric. For. Meteorol. 217, 89–100.
Griffin, R. C., Montgomery, J. M., Rister, M. E., 1987. Selecting functional form in production function analysis. Western J. Agr. Econ. 12 (2), 216–227.
Hansen, L., 1991. Farmer response to changes in climate the case of corn production. J. Agr. Econ. Res. 43 (4), 18–25.
Haude, W., 1955. Zur Bestimmung der Verdunstung auf möglichst einfache Weise. Vol. 11 of Mitteilungen des Deutschen Wetterdienstes. Deutscher Wetterdienst.
Heimfarth, L., Finger, R., Musshoff, O., 2012. Hedging weather risk on aggregated and individual farm-level: Pitfalls of aggregation biases on the evaluation of weather index-based insurance. Agr. Finance Rev. 72 (3), 471–487.
Hussain, S. S., Mudasser, M., 2007. Prospects for wheat production under changing climate in mountain areas of Pakistan – an econometric analysis. Agric. Syst. 94 (2), 494–501.
IEEP, 2006. Implementation of articles 18, 19, 20 and 16 of regulation (ec) no. 1257/1999 in the 25 member states of the European Union. report. Institute of European Environmental Policy, accessed 04.02.2014, available at http://ec. europa.eu/agriculture/eval/reports/lfa/full_annex_en.pdf.
Isik, M., Devadoss, S., 2006. An analysis of the impact of climate change on crop yields and yield variability. Appl. Econ. 38 (7), 835–844. U
Kaufmann, R. K., Snell, S. E., 1997. A biophysical model of corn yield: Integrating climatic and social determinants. Amer. J. Agr. Econ. 79 (1), 178–190.
Kaylen, M. S., Wade, J. W., Frank, D. B., 1992. Stochastic trend, weather and US corn yield variability. Appl. Econ. 24 (5), 513–518.
Li, S., Wheeler, T., Challinor, A., Lin, E., Ju, H., Xu, Y., 2010. The observed relationships between wheat and climate in China. Agric. For. Meteorol. 150, 1412–1419.
Lobell, D. B., 2007. Changes in diurnal temperature range and national cereal yields. Agric. For. Meteorol. 145 (3-4), 229–238.
Lobell, D. B., Asner, G. P., 2003. Climate and management contributions to recent trends in U.S. agricultural yields. Science 299, 1032.
Lobell, D. B., Burke, M. B., Tebaldi, C., Mastrandrea, M. D., Falcon, W. P., Naylor, R. L., 2008. Prioritizing climate change adaptation needs for food security in 2030. Science 319 (5863), 607–610.
Lobell, D. B., Ortiz-Monasterio, J. I., Asner, G. P., Matson, P. A., Naylor, R. L., Falcon, W. P., 2005. Analysis of wheat yield and climatic trends in Mexico. Field Crop. Res. 2, 250–256.
Lobell, D. B., Roberts, M. J., Schlenker, W., Braun, N., Little, B. B., Rejesus, R. M., Hammer, G. L., 2014. Greater sensitivity to drought accompanies maize yield increase in the U.S. midwest. Science 344 (6183), 516–519.
Lobell, D. B., Schlenker, W., Costa-Roberts, J., 2011. Climate trends and global crop production since 1980. Science 333 (6042), 616–620.
Miao, R., Khanna, M., Huang, H., 2016. Responsiveness of crop yield and acreage to prices and climate. Amer. J. Agr. Econ. 98 (1), 191–211.
Nicholls, N., May 1997. Increased Australian wheat yield due to recent climate trends. Nature 387, 484–485.
Odening, M., Musshoff, O., Xu, W., 2007. Analysis of rainfall derivatives using daily precipitation models: Opportunities and pitfalls. Agr. Finance Rev. 67 (1), 135– 156.
Ortiz-Bobea, A., Just, R. E., 2013. Modeling the structure of adaptation in climate change impact assessment. Amer. J. Agr. Econ. 95 (2), 244–251.

Osborne, T. M., Wheeler, T. R., 2013. Evidence for a climate signal in trends of global crop yield variability over the past 50 years. Environ. Res. Lett. 8 (2), 1–9.
R Development Core Team, 2016. R: A Language and Environment for Statistical Computing. Vienna, accessed 18.05.2016, avaiable at http://www.R-project. org/.
Ray, D. K., Gerber, J. S., MacDonald, G. K., West, P. C., 2015. Climate variation explains a third of global crop yield variability. Nat. Commun. 6 (5989), 1–9.
Reidsma, P., Ewert, F., Lansink, A. O., 2007. Analysis of farm performance in Europe under different climatic and management conditions to improve understanding of adaptive capacity. Clim. Chang. 84 (3-4), 403–422.
Roberts, M. J., Schlenker, W., Eyer, J., 2013. Agronomic weather measures in econometric models of crop yield with implications for climate change. Amer. J. Agr. Econ. 95 (2), 236–243.
Schlenker, W., Lobell, D. B., 2010. Robust negative impacts of climate change on African agriculture. Environ. Res. Lett. 5 (1), 1–8.
Schlenker, W., Roberts, M. J., 2009. Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. Proc. Natl. Acad. Sci. 106 (37), 15594–15598.
Schrödter, H., 1985. Verdunstung: Anwendungsorientierte Meßverfahren und Bestimmungsmethoden, 1st Edition. Springer, Berlin.
Sonntag, D., 1990. Important new values of the physical constants of 1986, vapour pressure formulations based on its-90, and psychrometer formulae. Meteor. Z. 4 (5), 340–344.
Statistisches Bundesamt, 2014a. Erntemenge (Feldfrüchte und Grünland): Deutschland, Jahre, Fruchtarten, Tabellencode: 41241-0005. Accessed 11.11.2014, available at https://www-genesis.destatis.de/.
Statistisches Bundesamt, 2014b. Index der Einkaufspreise landwirtschaftl. Betriebsmittel: Deutschland, Wirtschaftsjahr, Messzahlen mit/ohne Umsatzsteuer, Landwirtschaftliche Betriebsmittel. Tabellencode: 61221-0002. Accessed 29.10.2014, available at https://www-genesis.destatis.de/genesis/online.
Statistisches Bundesamt, 2014c. Preise Preisindizes für die Land- und Forstwirtschaft. No. 11/2013 in Fachserie 17 Reihe 1. Statistisches Bundesamt, Wiesbaden, accessed 12.02.2014, available at https://www.destatis.de/DE/Publikationen/Thematisch/Preise/Landwirtschaftspreise/ErzeugerpreiseLandForstwirtschaft2170100131114.pdf? blob= publicationFile.
Ward, P. S., Florax, R. J. G. M., Flores-Lagunes, A., 2014. Climate change and agricultural productivity in Sub-Saharan Africa: A spatial sample selection model. Europ. Rev. Agr. Econ. 41 (2), 199–226.
Wooldridge, J. M., 2009. Introductory Econometrics. A Modern Approach, 4th Edition. South-Western Cengage Learning, Mason, Ohio.
Yang, S.-R., Koo, W. W., Wilson, W. W., 1992. Heteroskedasticity in crop yield models. J. Agr. Resource Econ. 17 (1), 103–109.
You, L., Rosegrant, M. W., Wood, S., Sun, D., 2009. Impact of growing season temperature on wheat productivity in China. Agric. For. Meteorol. 149, 1009–1014.

11.2 Level normalized modeling approach of yield volatility for winter wheat and silage maize on different scales within Germany

Christoph Gornott[1*] and Frank Wechsung[1]

[1] Potsdam Institute for Climate Impact Research (PIK)

[*] Corresponding author

11.2.1 Daten und Aggregation

Die Winterweizen- und Silomaiserträge liegen auf Landkreisebene für den Zeitraum von 1991 bis 2010 vor. Die Erträge von 1991 bis 1998 sind von uns aus den statistischen Jahrbüchern der deutschen Bundesländer digitalisiert worden. Die Erträge von 1999 bis 2010 sind über Statistische Ämter des Bundes und der Länder (2013b) digital verfügbar. In den Bundesländern Sachsen geht die Zeitreihe insgesamt nur von 1992 bis 2007, in Sachsen-Anhalt von 1991 bis 2006. Landkreise ohne oder mit unvollständigen Ertragsdaten bleiben unberücksichtigt. Der Wetterdatensatz enthält Temperatur als Tages-Maximum (T_{max}), -Minimum (T_{min}) und -Mittel (T_{avg}) sowie Globalstrahlung (R_S) und Niederschlag als Tagessummen (DWD, 2011). Räumlich wurden die 1218 Wetterstationen des Deutschen Wetterdienstes den Landkreisen zugewiesen. Bei mehreren Wetterstationen in einem Landkreis wurde das arithmetische Mittel genommen. Landkreise ohne Wetterstation und Wetterstationen oberhalb von 700 m über Normalhöhennull blieben unberücksichtigt. Wir verwendeten diese Höhenrestriktion, da in Deutschland oberhalb davon kein Ackerbau praktiziert wird. Die ökonomischen Proxyvariablen Anbaufläche und Düngerpreis liegen auf Deutschlandebene vor. Die Anbaufläche von Weizen und Silomais wurde aus den Datensätzen der Statistischen Ämter des Bundes und der Länder (2013a) [1991 bis 2008] und des Statistisches Bundesamt (2013) [2008 bis 2010] zusammengesetzt. Der Düngerpreisindex und weitere getestete Faktor- und Produktpreise kommen von den Statistical offices of the Federation and the Länder (2013) und der Statistische Ämter des Bundes und der Länder (2013c).

11.2.2 Software

Die Modelle wurden mit der Software R (R Core Team, 2013) geschätzt. Für die PDMs nutzten wir das R-Software-Paket *plm* (Croissant und Millo, 2008), für die RCM das *lme4*-Paket (Bates, 2010) und für die statistischen Tests das *lmtest*-Paket (Zeileis und Hothorn, 2002). Die robusten Standardfehler nach Arellano wurden über das *sandwich*-Paket errechnet (Zeilleis, 2004). Die Zuweisung der Wetterstationen und die Aggregation zu (Sub)Nationen erfolgten über das *sqldf*-Paket (Grothendieck, 2012). Die Karten wurden mit dem grafischen Informationssystem Q-GIS erstellt.

11.2.3 Errechnung der Wachstumsgradtage

Die nachfolgende Formel zeigt die Berechnung der Wachstumsgradtage (WGT) aus T_{UL} mit 8°C und T_{OL} mit 32°C als unsteteres und oberes Limit (Roberts et al., 2012):

$$WGT = \begin{cases} T_{OL} - T_{UL} & \text{wenn} \quad T_{avg} > T_{OL} \\ T_{avg} - T_{UL} & \text{wenn} \quad T_{avg} \leq T_{OL}. \\ 0 & \text{wenn} \quad T_{avg} \leq T_{UL} \end{cases}$$

11.2.4 Nicht signifikanter Variablen

Statistisch signifikante Parameter sind nicht alleine für die Variablenauswahl ausschlaggebend (Wooldridge, 2013: 127-129). Nuzzo (2014) zeigt, dass durch den *p*-Wert Ergebnisse plausibler gemacht werden können. Eine hohe statistische Signifikanz bedeutet aber nur, dass die Wahrscheinlichkeit des richtigen Ergebnisses steigt. Wooldridge (2013: 141) beschreibt, dass einzeln statistisch signifikante Variablen in der Kombination mit anderen Variablen nicht mehr signifikant sind und *vice versa*. Nach Studenmund (2000: 172-173) ist eine schrittweise Regression, die nacheinander signifikante Variablen in ein Modell aufnimmt, ein Eingeständnis von Unwissenheit über die Variablenauswahl. Wegen der willkürlichen Reihenfolge mit der die Auswahl erfolgt, gibt es keine theoretische (kausale) Begründung für die Variablenauswahl.

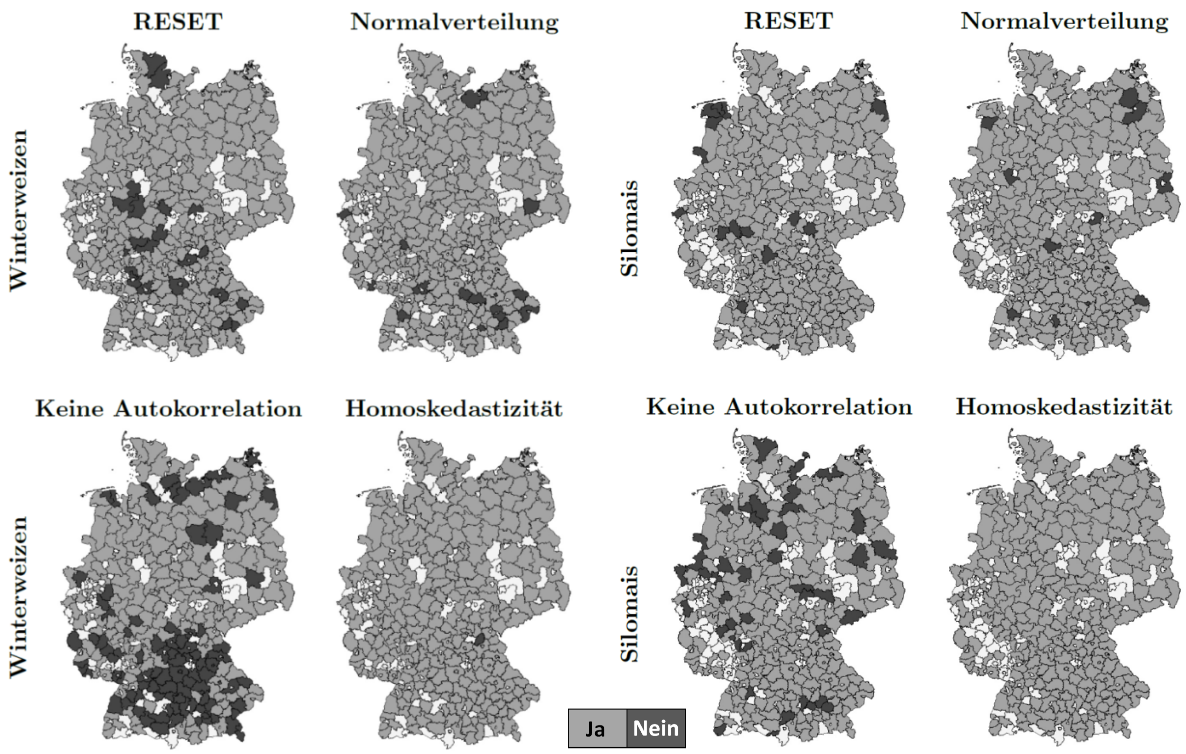


Abb. A1. Statistische Tests der STSMs für Winterweizen und Silomais: Funktionale Form (RESET): Nicht Fehlspezifiziert J/N, Normalverteilt (Shapiro-Wilk-Test): J/N, Keine Autokorrelation (Breusch-Godfrey/ Wooldridge-Test): J/N, Homoskedastisch (Breusch-Pagan-Test): J/N

11.2.5 Literatur

Bates, D., 2010: lme4: Mixed-effects modeling with R. Springer.
Croissant, Y., G. Millo, 2008: Panel Data Econometrics in R: The plm Package. Journal of Statistical Software, **27** (2), 1-43.
DWD, 2011: Tägliche Wetterdaten, 1951-2010. Deutscher Wetterdienst.
Grothendieck, G., 2012: sqldf: Perform SQL selects on R data frames. R-Package, 0.4-6.4.
Nuzzo, R., 2014: Scientific method: Statistical errors - P values, the 'gold standard' of statistical validity, are not as reliable as many scientists assume. Nature Climate Change, 506, 150-152.
R Core Team, 2013: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, 3.0.0. R Foundation for Statistical Computing, Vienna, Austria.
Roberts, M.J., W. Schlenker, J. Eyer, 2012: Agronomic weather measures in econometric models of crop yield with implications for climate change. American Journal of Agricultural Economics, 95 (2), 236-243.
Statistical offices of the Federation and the Länder, 2013: Index der Einkaufspreise landwirtschaftlicher Betriebsmittel, Index der Erzeugerpreise landwirtschaftlicher Produkte.
Statistische Ämter des Bundes und der Länder, 2013a: Datensatz Anbaufläche (Feldfrüchte und Grünland): Deutschland, Jahre, Fruchtarten 1991-2007.
Statistische Ämter des Bundes und der Länder, 2013b: Hektarerträge ausgewählter landwirtschaftlicher Feldfrüchte - Jahressumme - regionale Tiefe: Kreise und krfr. Städte.
Statistische Ämter des Bundes und der Länder, 2013c: Index der Erzeugerpreise landwirtschaftlicher Produkte.
Statistisches Bundesamt, 2013: Ackerland nach Hauptfruchtgruppen und Fruchtarten, Land- & Forstwirtschaft, Fischerei - Feldfrüchte und Grünland.
Studenmund , A.H., 2000: Using Econometrics: A Practical Guide, 4. Addison Wesley.
Wooldridge, J.M., 2013: Introductory Econometrics. A Modern Approach. South Western Cengage Learning, 868 S.
Zeileis, A., T. Hothorn, 2002: Diagnostic Checking in Regression Relationships.
Zeilleis, A., 2004: Econometric Computing with HC and HAC Covariance Matrix Estimators. Journal of Statistical Software, 11 (10), 1-17.

11.3 Statistical regression models for assessing climate impacts on crop yields: A validation study for winter wheat and silage maize in Germany

Christoph Gornott[1*] and Frank Wechsung[1]

[1] Potsdam Institute for Climate Impact Research (PIK)

[*] Corresponding author

11.3.1 Data and aggregation

Winter wheat and silage maize yield data are available on county scale for the period from 1991 to 2010. The yields between 1991 and 1998 are digitized from the statistical yearbooks of the German federal states. The yields between 1999 and 2010 are digitally available from the Statistical Offices of the Federation and the Länder (2013b). The time series for Saxony last only from 1992 to 2007 and those for Saxony-Anhalt from 1991 to 2006. Counties without or with incomplete yield data are not considered in our analysis. The weather data contains temperature as daily maximum ($T_{\max}$), minimum ($T_{\min}$), and average (T_{avg}) and solar radiation (R_S) and precipitation as daily sums. The weather data are based on measurements at 1,218 weather stations of the German weather service (DWD, 2011) within Germany. The weather stations are assigned to the counties according to their location. In case of more than one weather station per county, we take the arithmetic average of all stations. Counties without weather stations and the weather stations above an altitude of 700m are unconsidered (6.9% of the 1,218 weather stations). This altitude restriction is chosen because husbandry is not practiced above this altitude in Germany. The economic proxy variables acreage and fertilizer price are observed only on national scale (for Germany). The acreage data of winter wheat and silage maize is based on datasets of the Statistical Offices of the Federation and the Länder (2013a) [1991 to 2008] and the Federal Statistical Office (2013) [2008 to 2010]. The fertilizer price (and further factor and product prices) is from the Statistical Offices of the Federation and the Länder (2013c).

11.3.2 Using statistically not significant variables

Statistical significance is not the only criteria for the variable selection (Wooldridge, 2013, p. 127-129). Nuzzo (2014) shows that results can be made also more plausible by a low p-value. A high statistical significance means that the probability of the correct result increases. Wooldridge (2013, p. 141) describes that individually statistically significant variables in combination with other variables are often no longer significant and *vice versa*. Studenmund (2000, p. 172-173) criticizes that a stepwise regression, which takes successively significant variables in a model, is “an admission of ignorance” of the variable selection. The arbitrary order to select variables prevents a plant-physiologically reasonable selection of the variables.

Prost et al. (2008) and Whittingham et al. (2006) show the limitations of a stepwise regression, because of the variable section can be biased due to the selection procedure and the selection criteria. Furthermore, the variable selection highly depends on the estimation dataset and is only limited exportable to other datasets, regions ore time periods. Finally, important variables, like the precipitation, are occasionally not considered by the stepwise approach. In such a case, the projections might be affected by an omitted variable bias.

11.3.3 Model fit

The NSE (Eq. S.1), the R^2 (Eq. S.2), and the RMSE (Eq. S.3) are calculated by the estimated (E) and observed (O) yield changes by the following equations (the bar means the arithmetic average):

$$NSE = 1 - \frac{\sum_{t=1}^{M}(O-E)^2}{\sum_{t=1}^{M}(O-\bar{O})^2}, \qquad \text{with } t = 1, \dots, M \tag{S.1}$$

$$R^2 = \left(\frac{\sum_{t=1}^{M}(O-\bar{O})\,(E-\bar{E})}{\sqrt{\sum_{t=1}^{M}(O-\bar{O})^2}\,\sqrt{\sum_{t=1}^{M}(E-\bar{E})^2}}\right)^2 \tag{S.2}$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{M}(O-E)^2}{M}} \tag{S.3}$$

11.3.4 Software

The models are estimated utilizing the software *R* (R Core Team, 2013). We use the package *plm* for the PDMs (Croissant and Millo, 2008), the package *lme4* for the RCMs (Bates, 2010) and the package *lmtest* for the statistical tests (Zeileis and Hothorn, 2002). The robust standard errors after Arellano are computed using the *sandwich* package (Zeileis, 2004). The assignment of the weather stations and the aggregation of counties to (sub)-nations, we carried out using the *sqldf* package (Grothendieck, 2012). The maps are generated with the geographic information system software *Arc-GIS*.

11.3.5 Statistical tests

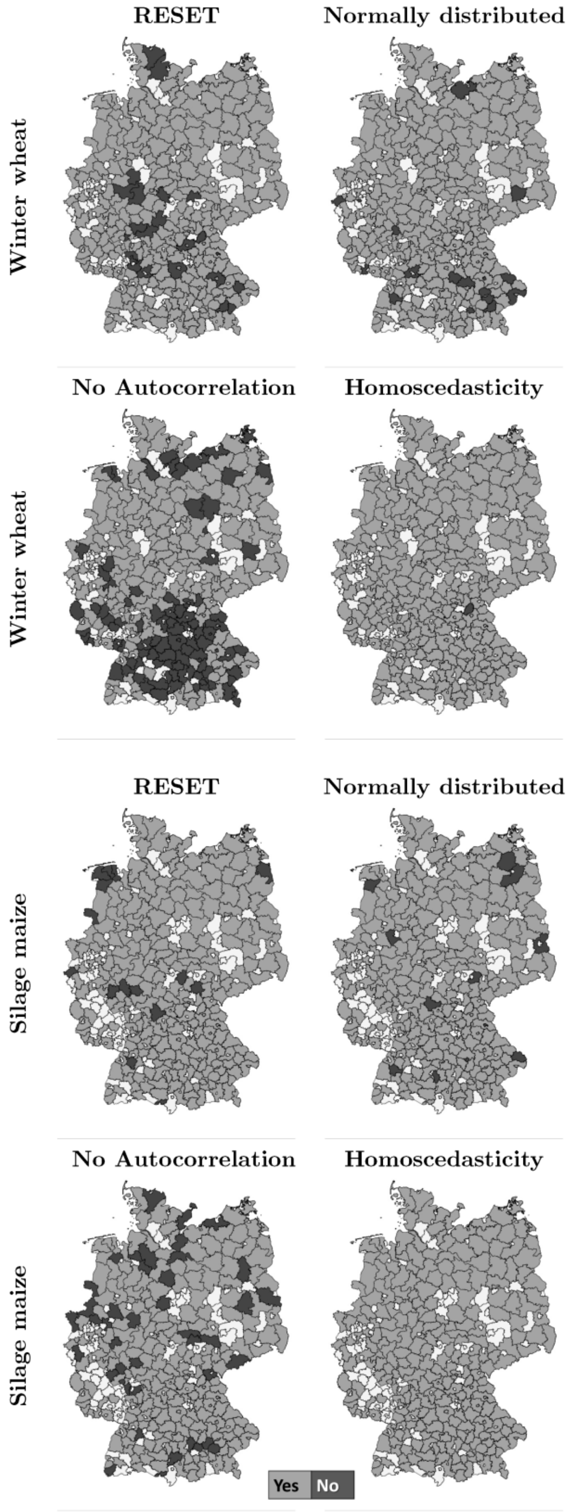


Fig. S.1: Statistical Tests of the STSMs: Regression Equation Specification Error Test (RESET, functional form): not error specified: Yes/ No; normal distributed (Shapiro-Wilk-Test): Yes/ No; no autocorrelation (Breusch-Godfrey/ Wooldridge-test): Yes/ No; homoscedasticity (Breusch-Pagan test): Yes/ No. The statistical tests are carried out with the R package *lmtest* (Zeileis and Hothorn, 2002).

Tab. S.1: Statistical Tests of the PDMs: The following statistical tests are binary coded p ≤ 0.01 → ✓, *p* > 0.01 → ×: FF (functional form): RESET Regression Equation Specification Error Test (RESET), ✓ = error specified, squared components have a significant effect; LM: Lagrange-Multiplier-test, ✓ = significant differences across counties; BG: Breusch-Godfrey/ Wooldridge-test ✓ = autocorrelation; BP: Breusch-Pagan test of heteroscedasticity, ✓ = heteroscedasticity. The statistical tests are carried out with the R package lmtest (Zeileis and Hothorn, 2002).

(Sub)Nation	FF	LM	BG	BP
Winter wheat				
Schleswig-Holstein	×	✓	✓	✓
Lower Saxony	×	✓	✓	✓
North Rhine-Westphalia	✓	✓	✓	✓
Hesse	×	✓	✓	✓
Rhineland-Palatinate	×	✓	✓	✓
Baden-Württemberg	×	✓	✓	×
Bavaria	✓	✓	✓	×
Saarland	×	×	✓	×
Brandenburg	✓	✓	✓	×
Mecklenburg-W. Pomerania	✓	✓	✓	✓
Saxony	×	✓	✓	×
Saxony-Anhalt	×	✓	✓	✓
Thuringia	✓	✓	✓	✓
Schlei/ Trave	×	✓	✓	×
Elbe	✓	✓	✓	✓
Weser	×	✓	✓	✓
Ems	×	✓	✓	×
Rhine	×	✓	✓	✓
Maas	✓	✓	×	×
Danube	×	✓	✓	×
Warnow/ Peene	✓	✓	✓	✓
Oder	×	✓	×	×
Germany	✓	✓	✓	✓
Silage maize				
Schleswig-Holstein	×	✓	✓	×
Lower Saxony	×	✓	✓	×
North Rhine-Westphalia	×	✓	✓	×
Hesse	×	✓	✓	×
Rhineland-Palatinate	✓	✓	✓	✓
Baden-Württemberg	✓	✓	✓	×
Bavaria	✓	✓	✓	✓
Saarland	✓	×	×	✓
Brandenburg	✓	✓	✓	×
Mecklenburg-W. Pomerania	✓	✓	✓	×
Sachsen	✓	×	✓	×
Saxony-Anhalt	✓	✓	✓	×
Thuringia	✓	✓	✓	×
Schlei/ Trave	×	✓	✓	×
Elbe	✓	✓	✓	✓
Weser	×	✓	✓	✓
Ems	×	✓	✓	×
Rhine	×	✓	✓	×
Maas	×	✓	×	×
Danube	✓	✓	✓	×
Warnow/ Peene	✓	✓	✓	×
Oder	×	✓	×	×
Germany	✓	✓	✓	✓

Tab. S.2: Correlation (Pearson) coefficients of the variables. The acronyms are: precipitation – PREC, potential evapotranspiration – ETP, temperature normalized solar radiation – SRT, solar radiation – R_S, fertilizer price – Fert, and acreage winter wheat – Ac WW. The month behind the variables are the corresponding period.

Variable	PREC	ETP	PREC	ETP	PREC	ETP	R_S	SRT	Fert	Ac WW
Period	May–Jul		Nov–Apr		Aug–Oct		May–Jul			
PREC May–Jul	1.00									
ETP May–Jul	-0.13	1.00								
PREC Nov–Apr	0.25	0.05	1.00							
ETP Nov–Apr	0.38	0.44	-0.04	1.00						
PREC Aug–Oct	0.34	-0.06	0.45	0.03	1.00					
ETP Aug–Oct	-0.06	0.61	-0.09	0.46	-0.39	1.00				
R_S May–Jul	-0.20	0.48	0.10	-0.13	-0.04	0.19	1.00			
SRT May–Jul	-0.16	0.14	0.09	-0.31	-0.06	0.04	0.87	1.00		
Fert	0.04	0.18	-0.06	0.17	-0.01	-0.05	0.07	0.02	1.00	
Ac WW	0.03	0.23	-0.02	0.18	0.10	0.05	0.06	-0.03	0.72	1.00

11.3.6 Further description of the parameters

The SRT May-July north-south parameter gradient for wheat reflects a similar gradient of the absolute SRT levels. The sensitivity is often high (low) in regions with low (high) absolute values. However, this relationship also has exceptions. The sensitivity is high and the relationship directly proportional at the coast line where the absolute level of values there is high as well. The higher responsiveness of wheat to SRT compared with silage maize is also reflected in the spatial patterns. The north south decline is not only weaker for silage maize than for winter wheat, but the parameter values even reverse in the east. The reason for this is unclear. A higher frequency of late frost events, which could be related to increased SRT values, seems to be a reasonable speculation.

Fig. S.2 depicts the yield impact of inter-annual changes in precipitation ($PREC$) from May to July for winter wheat determined from STSMs. The x-axis gives values of the precipitation parameter (β_{PREC}) normalized to the bulk of other parameters β_j (Eq. S.4). To make the positive and negative parameter values comparable, we use the absolute parameter values (hereafter expressed as sensitivity portion of $PREC$ May-Jul). The y-axis shows explained yield variability of precipitation ($\beta_{PREC}\ \log x'_{PREC\ t}$) normalized to the total explained yield variability (Eq. S.5). This term is expressed as variance portion of $PREC$ May-Jul. The term var is the variance.

$$\text{sensitivity portion of } PREC \text{ May} - \text{Jul} = |\beta_{PREC}| \left(\sum_{j=1}^{J} |\beta_j| \right)^{-1} \tag{S.4}$$

$$\text{variance portion of } PREC \text{ May} - \text{Jul} = \text{var}\ (\beta_{PREC}\ \log x'_{PREC\ t}) \left(var \sum_{j=1}^{J} \beta_j\ \log x'_{jt} \right)^{-1} \tag{S.5}$$

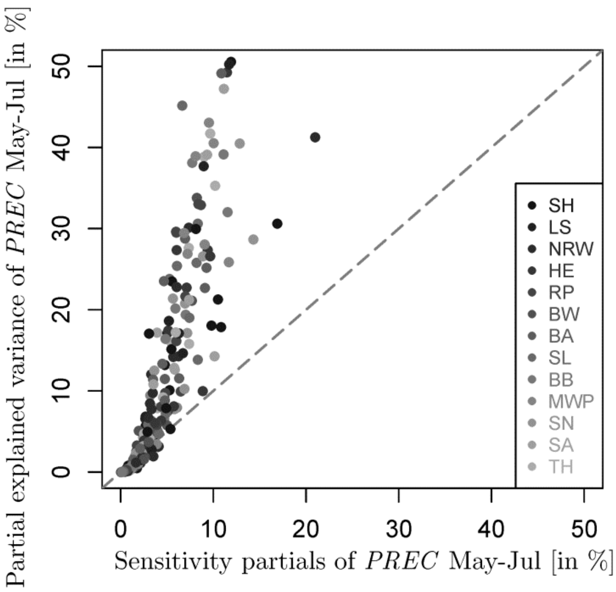


Fig. S.2: Partial explained variance of *PREC* May-Jul and sensitivity partials of *PREC* May-Jul for winter wheat determined for the STSMs. The acronyms are: SH: Schleswig-Holstein, LS: Lower Saxony, NRW: North Rhine-Westphalia, HE: Hesse, RP: Rhineland-Palatinate, BW: Baden-Württemberg, BA: Bavaria, SL: Saarland, BB: Brandenburg, MWP: Mecklenburg-Western Pomerania, SN: Saxony, SA: Saxony-Anhalt, TH: Thuringia.

11.3.7 References

Bates, D., 2010. lme4: Mixed-effects modeling with R. Springer.

Croissant, Y. and Millo, G., 2008. Panel data econometrics in R: The plm package. Journal of Statistical Software, 27(2): 1-43.

DWD, 2011. Daily weather data, 1951 - 2010. German Weather Service.

Federal Statistical Office, 2013. Ackerland nach Hauptfruchtgruppen und Fruchtarten.

Grothendieck, G., 2012. sqldf: Perform SQL selects on R data frames. R-Package, 0.4-6.4.

Nuzzo, R., 2014. Scientific method: Statistical errors - p values, the 'gold standard' of statistical validity, are not as reliable as many scientists assume. Nature Climate Change, 506: 150-152.

Prost, L., Makowski, D. and Jeuffroy, M.-H., 2008. Comparison of stepwise selection and Bayesian model averaging for yield gap analysis. Ecological Modelling, 219: 66-76.

R Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, 3.0.0. R Foundation for Statistical Computing, Vienna, Austria.

Schrödter, H., 1985. Verdunstung: Anwendungsorientierte Meßverfahren und Bestimmungsmethoden. Springer, 204 p.

Statistical Offices of the Federation and the Länder, 2013a. Datensatz Anbaufläche (Feldfrüchte und Grünland): Deutschland, Jahre, Fruchtarten 1991-2007.

Statistical Offices of the Federation and the Länder, 2013b. Hektarerträge ausgewählter landwirtschaftlicher Feldfrüchte - Jahressumme - regionale Tiefe: Kreise und krfr. Städte.

Statistical Offices of the Federation and the Länder, 2013c. Index der Einkaufspreise landwirtschaftlicher Betriebsmittel.

Studenmund, A.H., 2000. Using Econometrics: A Practical Guide, 4. Addison Wesley.

Whittingham, M.J., Stephens, P.A., Bradbury, R.B. and Freckleton, R.P., 2006. Why do we still use stepwise modelling in ecology and behaviour? Journal of Animal Ecology, 75: 1182–1189.

Wooldridge, J.M., 2013. Introductory Econometrics. A Modern Approach. South Western Cengage Learning, 868 pp.

Zeileis, A., 2004. Econometric Computing with HC and HAC Covariance Matrix Estimators. Journal of Statistical Software, 11(10): 1-17.

Zeileis, A. and Hothorn, T., 2002. Diagnostic Checking in Regression Relationships.

11.4 Global evaluation of a semiempirical model for yield anomalies and application to within-season yield forecasting

Bernhard Schauberger[1,2*], Christoph Gornott[1], Frank Wechsung[1]

[1] Potsdam Institute for Climate Impact Research (PIK)

[2] Laboratoire des Sciences du Climat et de l'Environnement, Institut Pierre-Simon Laplace (IPSL)

[*] Corresponding author

11.4.1 US climate regions, growing seasons, land-use patterns and reported yields used in this analysis

A map of the nine US climate regions used in this study is shown in Fig. S1. Their definition is based on Karl and Koss (1984). The average climate during the average maize growing season according to MIRCA2000 (Portmann et al., 2010) is shown in Fig. S2. Land use fractions according to MIRCA2000 are shown in Fig. S3. The distribution of growing seasons according to MIRCA2000 is shown in Fig. S4. For maize and soybeans calculated vegetative months of the average growing season in the US are April to June; calculated reproductive months are July to October. For spring wheat the split is May to June (vegetative) and July to August (reproductive). For winter wheat the vegetative part is October to April and the reproductive part May to July. The first months of the calculated reproductive seasons correspond with observed anthesis dates by the USDA[12]. Nationally aggregated yield time series, together with two anomaly calculation methods, are shown in Fig. S5. The equations used for aggregating grid cell time series to national averages are listed in supplementary equations S1 and S2. The equations used for defining PHUs and the ensuing split between vegetative and reproductive parts of the growing season are provided in equations S4 and S5 (section 2).

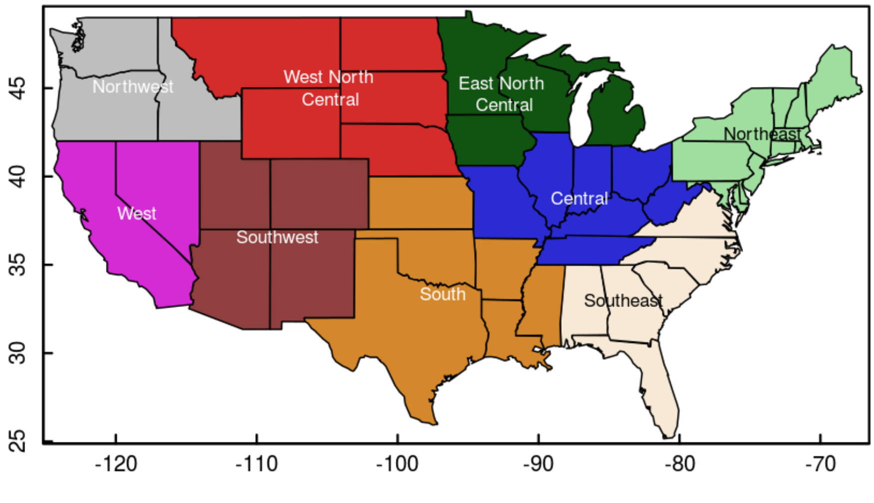


Fig. S1: The nine climate regions of the US as applied in this study.

[12] http://www.usda.gov/oce/weather/pubs/Other/MWCACP/MajorWorldCropAreas.pdf ; accessed on July 20, 2016

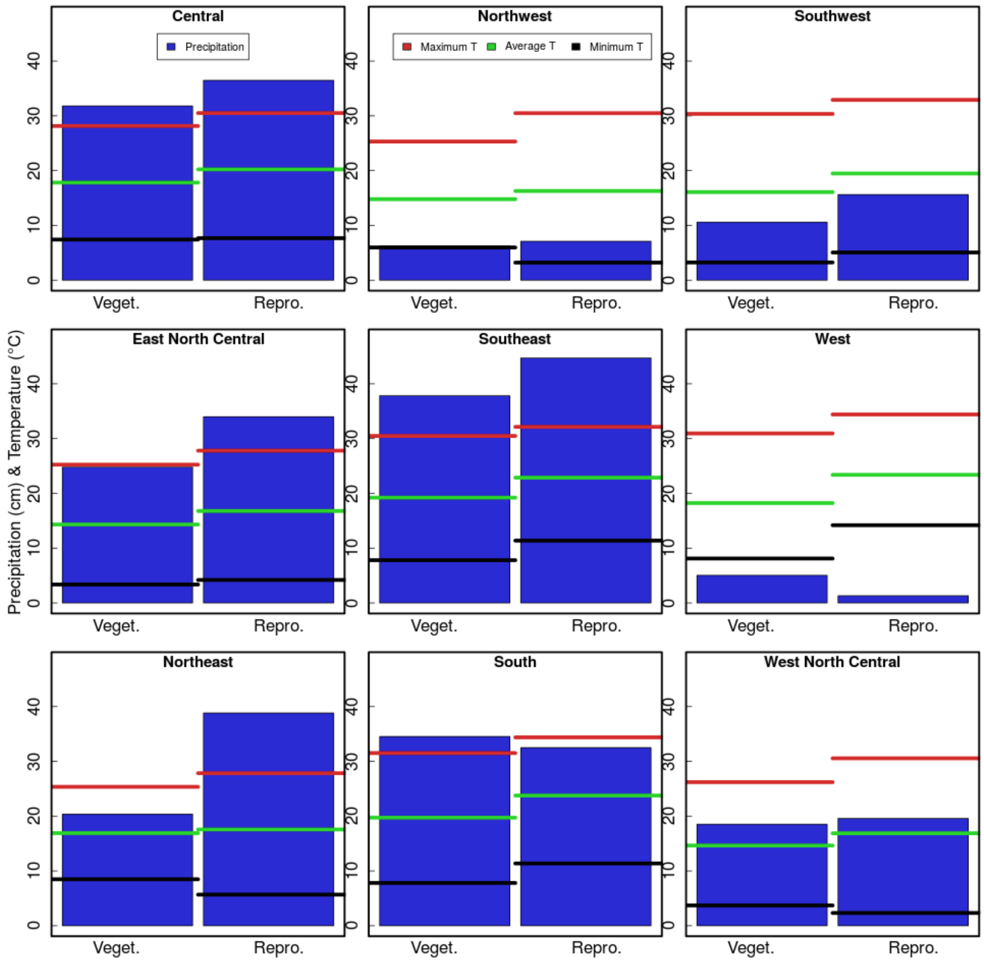
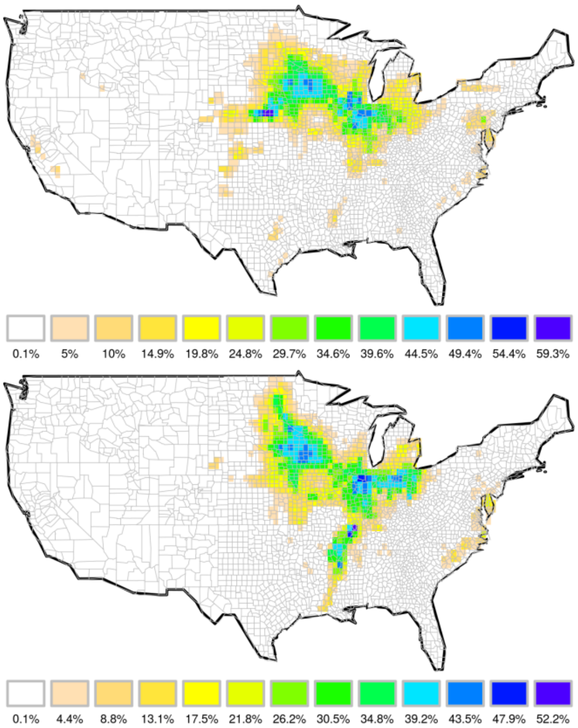
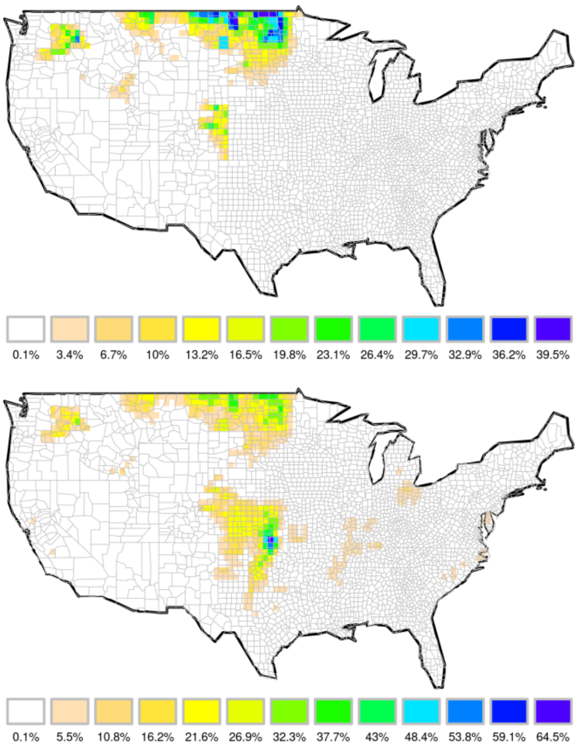


Fig. S2: Climate diagrams for the nine US climate regions during their maize growing season. Precipitation (in cm) and temperatures (in °C with minimum as blue, mean as green and maximum as red horizontal lines) are split into a vegetative and a reproductive part. Averages are calculated over space and time; the temperature extrema are averages over the individual grid cell extrema. The maize growing season according to MIRCA2000 (Portmann et al., 2010) can vary between regions (Fig. S4).

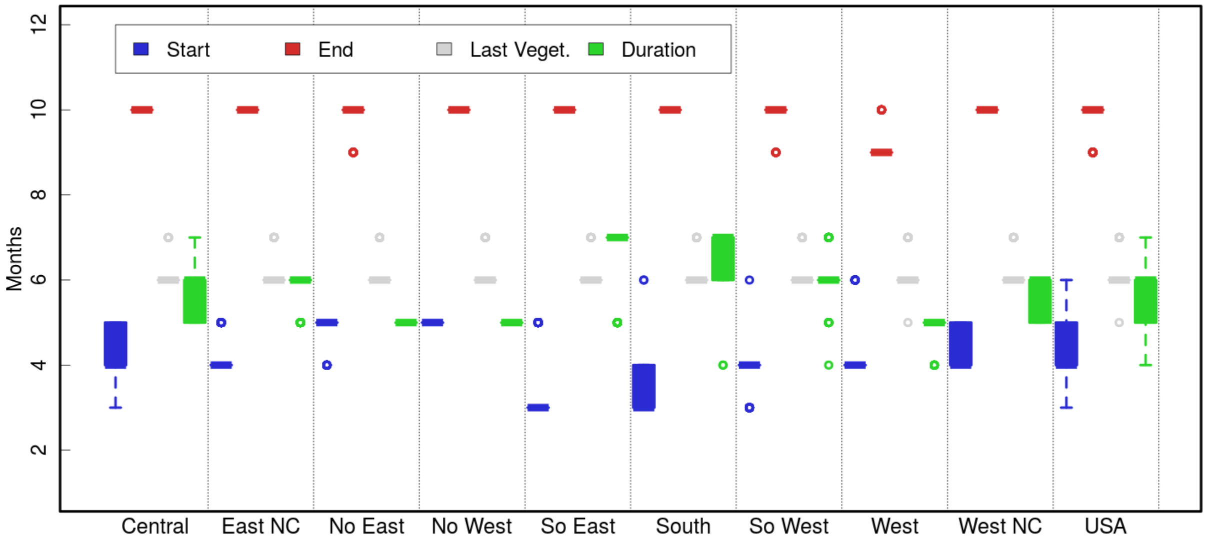


(a) Maize **(b)** Soybeans

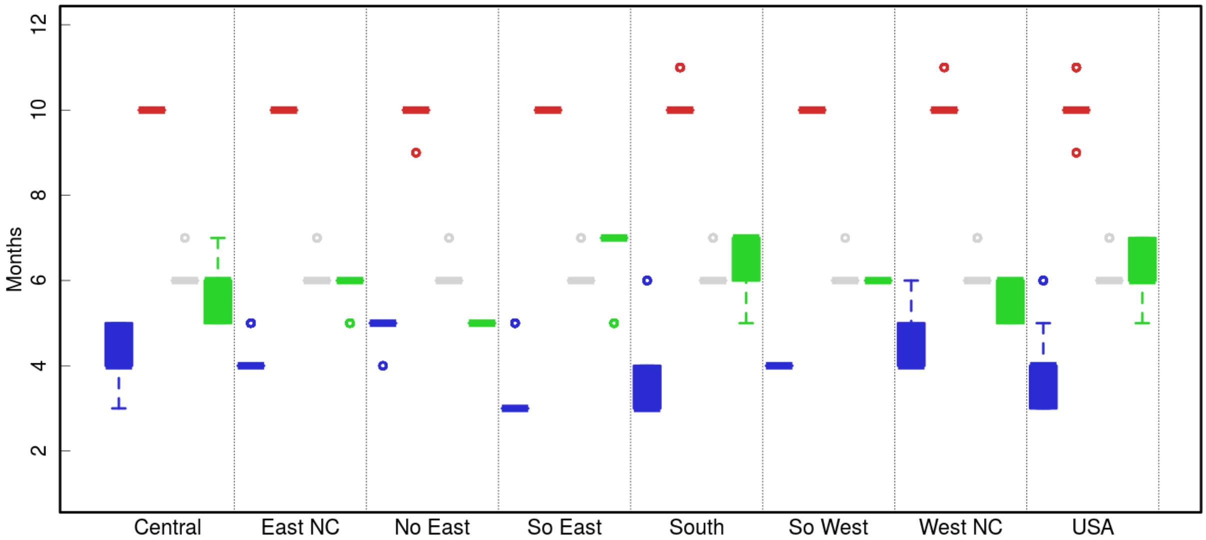


(c) Spring wheat **(d)** Winter wheat

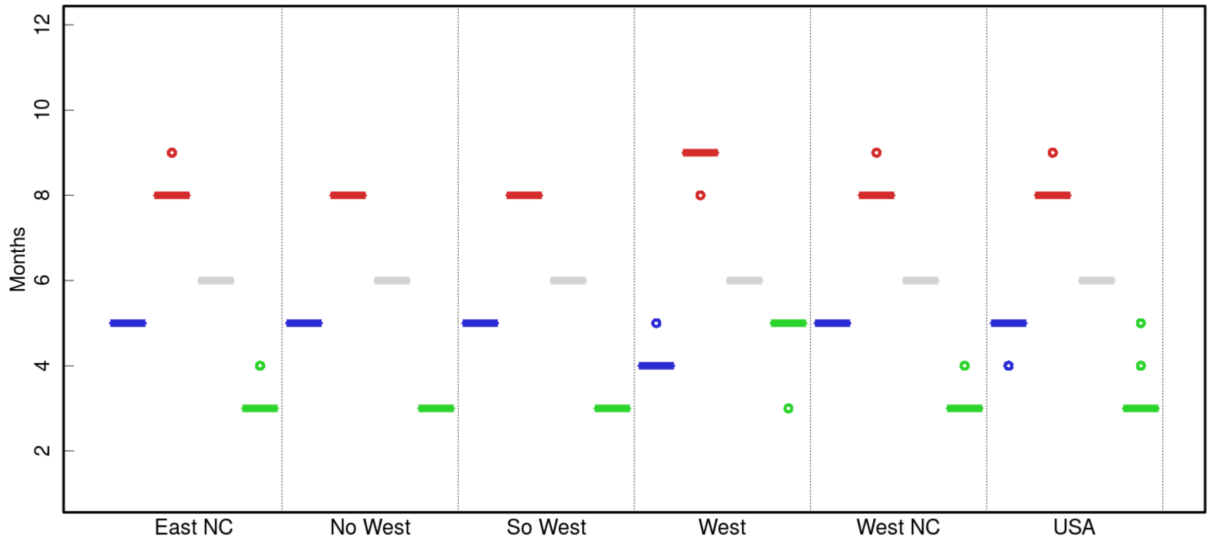
Fig. S3: Land use fractions as percent of grid cell area for maize (a), soybeans (b), spring wheat (c) and winter wheat (d) according to MIRCA2000 (Portmann et al., 2010). County boundaries are drawn in grey. Color scales diverge between land use maps.



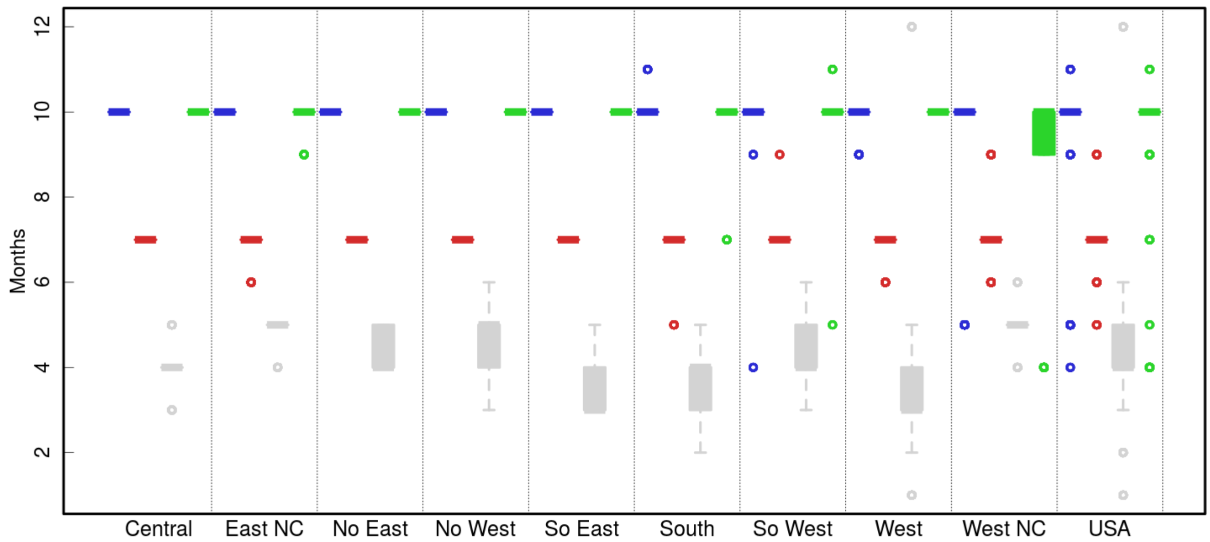
(a) Maize



(b) Soybeans



(c) Spring wheat



(d) Winter wheat

Fig. S4: Distribution of growing season start and end months (blue and red boxes, respectively), the last month of the vegetative growing season as defined by the 50% PHU threshold (see equations S4 and S5; grey boxes) and the duration of the growing season in months (green boxes). Several boxes are condensed to lines since there is no variation in the data. There is, in general, only little variation of growing seasons across the US according to MIRCA2000.

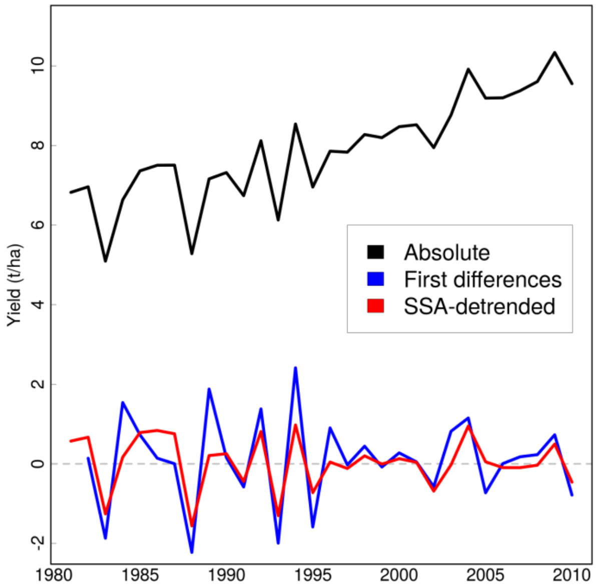
The equation used for aggregating time series from grid cells to climate region or country level is provided in equation S1.

$$\bar{y} = \frac{\sum_i l_i * a_i * y_i}{\sum_i l_i * a_i} \qquad \text{(S1)}$$

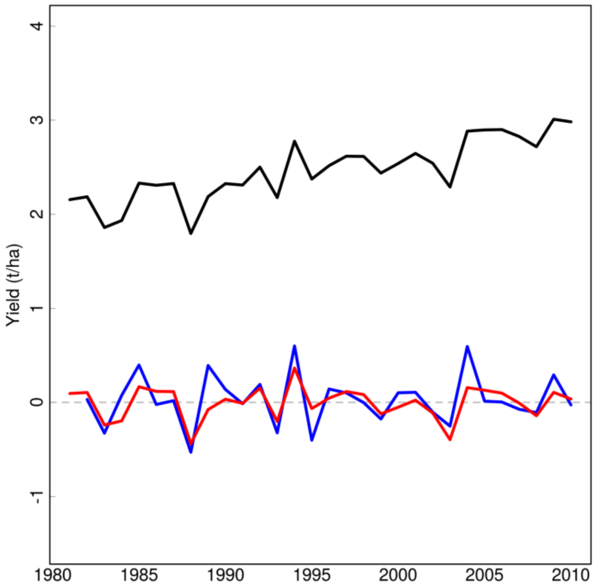
where y_i is yield anomaly in grid cell i, a_i is area of grid cell i, l_i is fraction of total land-use for the specific crop in grid cell i and $\bar{y}$ is the averaged yield anomaly over all grid cells in the aggregation region. If aggregation is weighted, l_i is taken from MIRCA2000 and a_i is calculated by equation S2. If aggregtion is unweighted, both l_i and a_i are set to 1, resulting in the standard average.

$$a_i = {r_E}^2 * \left(\lambda_{i,2} - \lambda_{i,1}\right) * (\sin\varphi_{i,2} - \sin\varphi_{i,1}) \qquad \text{(S2)}$$

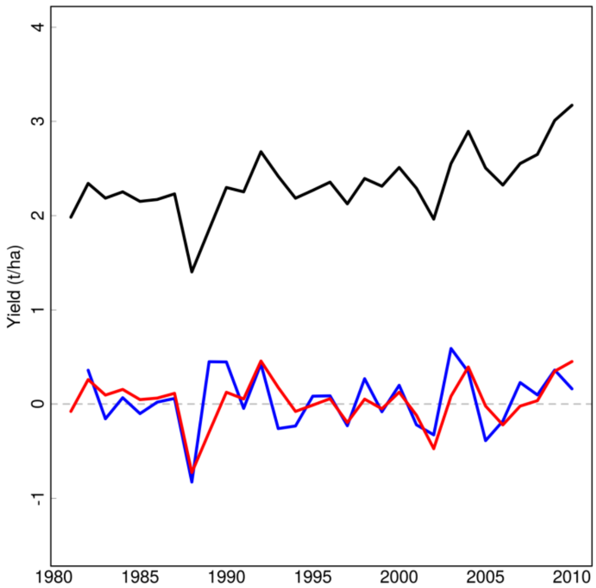
where r_E is earth radius (6,371 km), λ and φ are longitude and latitude boundaries of the grid cell (in radians).



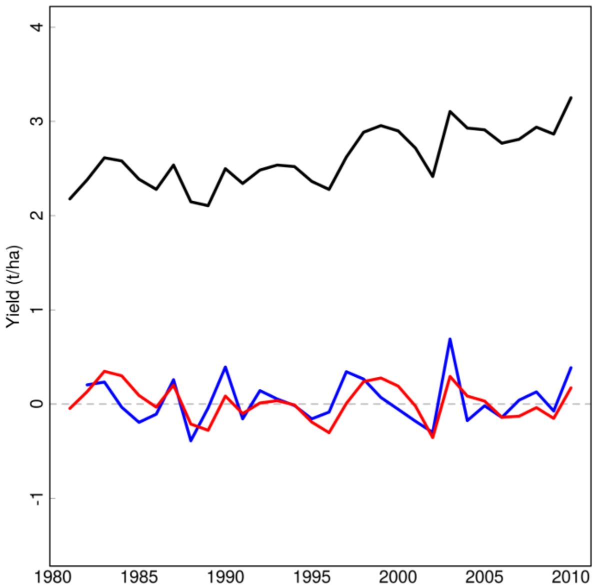
(a) Maize



(b) Soybeans



(c) Spring wheat



(d) Winter wheat

Fig. S5: Time series of yields for four staple US crops. Black lines are nationally aggregated yields, calculated from grid cells with land-use fraction for the respective crop larger than 0 and weighing by these land-use fractions. Blue and red lines, respectively, show yield anomalies calculated from these nationally aggregated yields by calculating either first differences between adjacent years (blue) or by subtracting a parameter-free Singular Spectrum Analysis trend (red).

11.4.2 Full regression equation

The full regression equation is provided in supplementary equation S3. This ‘standard’ regression can be modified by three switches (for sensitivity analyses): estimation method, included variables and anomaly calculation (Tab. S1). All combinations between all values (12 in total) are evaluated for each crop, yield data set and aggregation weighting combination (16 in total). This results in a total of 12 * 16 = 192 regressions for the US. The equations used to calculate phenological heat units and to deduce the transition month between vegetative and reproductive season are given in equations S4 and S5.

Equation S3 provides the fully specified 'standard' STSM regression formula for summer crops (i.e. with only the vegetative part of the temperature-corrected solar radiation). The equation contains eight coefficients including the intercept to be estimated ($\beta_{0\ldots7}$) for each grid cell. For PDMs fewer coefficients are estimated: there is only one set for $\beta_{0\ldots7}$ per aggregation region, but fixed effects allow for grid cell-specific intercepts.

$$\log Y_t' = \log\beta_0 + \beta_1 \log PET_{veg,t}' + \beta_2 \log PET'_{rep,t} + \beta_3 \log PR'_{veg,t} + \beta_4 \log PR'_{rep,t} + \beta_5 \log Rs'_{veg,t} + \beta_6 \log KDD_t' + \beta_7 \log FDD_t' + \log u_t'$$

(S3)

Variables are yield (*Y*), potential evapotranspiration (*PET*) during the vegetative (*veg*) and reproductive (*rep*) growing season parts, precipitation (*PR*) split into its vegetative and reproductive parts, temperature-corrected solar radiation (*SRT*) in the vegetative part of the growing season, killing and freezing degree days (*KDD*, *FDD*) over the whole growing season. The prime (′) behind each variable denotes yield anomalies. All variables are given for years (*t*) 1981 to 2010, starting one year later than data is available due to the first differences approach (two years later for winter wheat).

Tab. S1: Possible values for the three regression switches. All 12 combinations of the three specifiers are allowed.

Method	Variable set	Anomaly calculation
Separate Time Series Model ("STSM")	SRT (temperature-corrected solar radiation)	First differences ('first')
Panel Data Model ("PDM")	KDD-SRT (killing and freezing degree days plus SRT)	Difference to a singular spectrum analysis trend ('ssa')
	KDD-rad (KDD and FDD plus uncorrected solar radiation)	

Phenological heat units (PHU) above a base temperature over the growing season are calculated by equation S4:

$$PHU_d = \sum_{i=1}^{d} \max(T_i - T^{base}; 0) \qquad \text{(S4)}$$

where *d* is a day during the growing season (starting with 1), T_i is the temperature at day *i* and T^{base} is a crop-specific base temperature (8°C for maize and soybeans and 0°C for spring and winter wheat).

PHUs for a month are calculated by multiplying the PHU calculated from the monthly mean temperature by the number of days in this month.

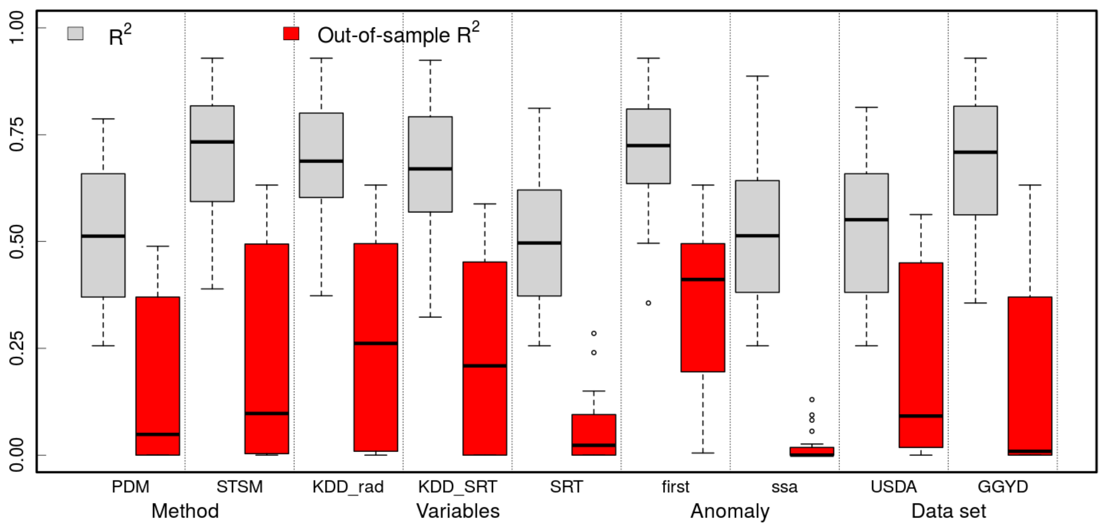
The first month of the reproductive season for each grid cell and crop is calculated by equation S5:

$$FRM = min_m(PHU_m \geq 0.5 * PHU_{max}) \qquad \text{(S5)}$$

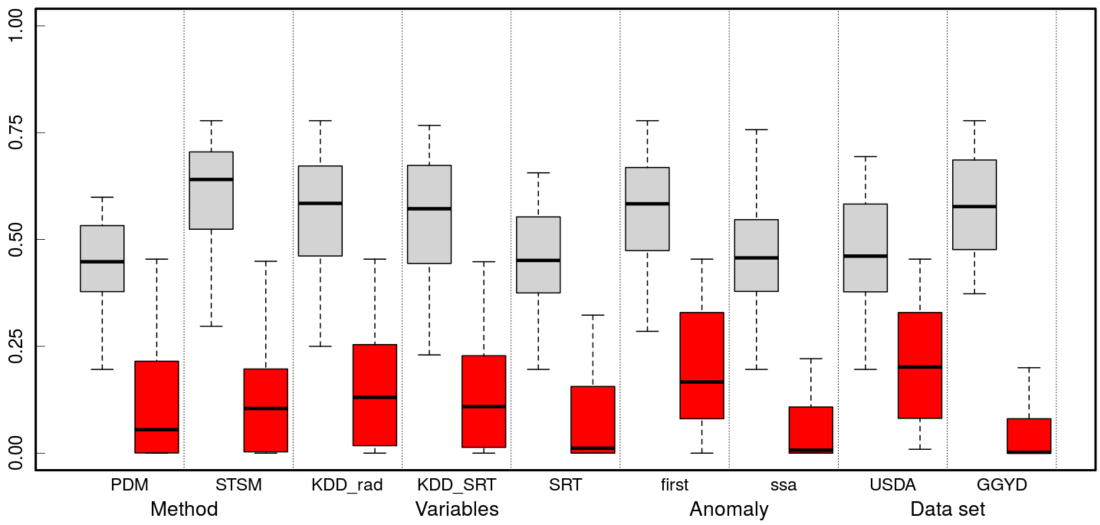
where *FRM* is "first reproductive month", *m* is a month in the growing season, PHU_m is the PHU for this month according to equation S4 and PHU_{max} is the total PHU achieved over the growing season.

11.4.3 Model evaluation in the US

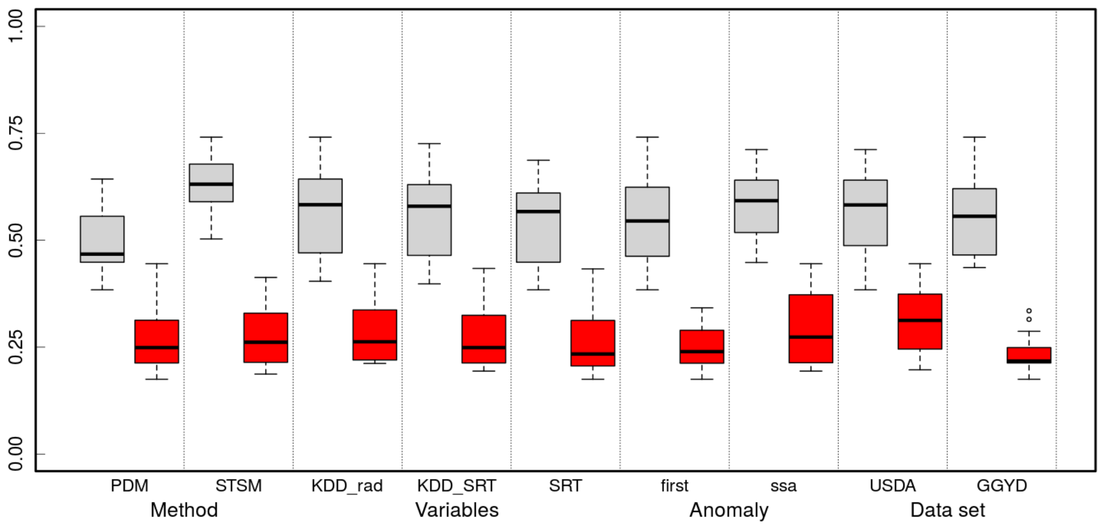
The performance ranges of different regression setups in the US are summarized in Fig. S6. For each crop the distributions of R^2 and R^2_{O1} across several model specifications are provided. Abbreviations are as follows. Regression method is either PDM (Panel Data Model) or STSM (Separate Time Series Model). Variables are either "KDD_rad" (uncorrected radiation instead of SRT, with KDD=Killing Degree Days and FDD=Freezing Degree Days variables), "KDD_SRT" (temperature-corrected radiation and KDD/FDD variables) or "SRT" (only temperature-corrected radiation, but without KDD/FDD). Yield anomaly calculation is done by either first differences ("first") or with a parameter-free trend estimated with Singular Spectrum Analysis ("ssa"). The data set can be either "USDA" (reported yield data provided by the USDA) or "GGYD" (global yield data derived from remote sensing and (sub)national yield statistics). There is a strong discrepancy between regression coefficients estimated from either reported or GGYD yield data, for both STSM (Fig. S7) and PDM (Fig. S8) estimation.



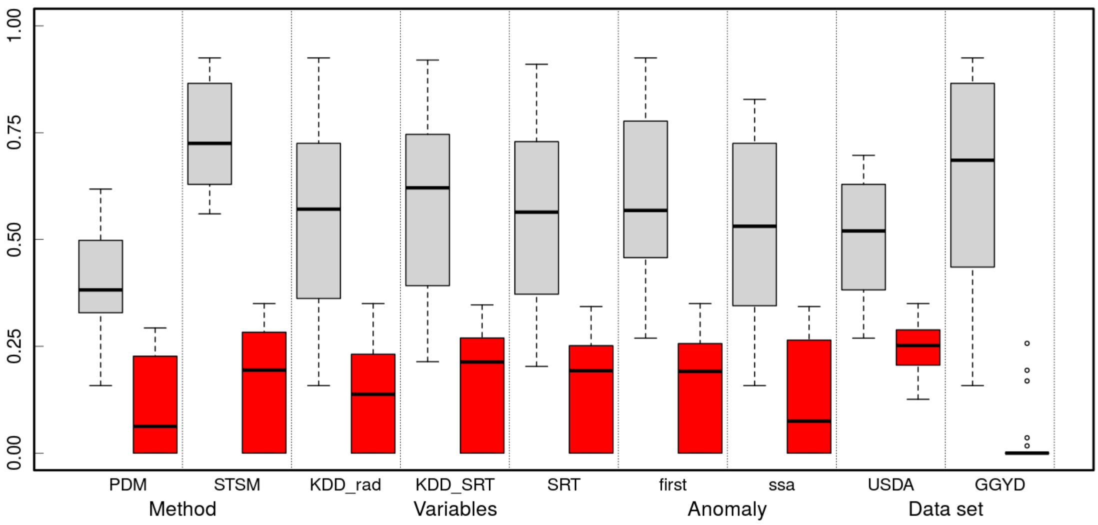
(a) Maize



(b) Soybeans



(c) Spring wheat



(d) Winter wheat

Fig. S6: Mean model performance (R^2 in gray and R^2_{01} in red) in the US for different model specifications, split by crops. Abbreviations of the model specifications are provided in the text and Tab. S1.

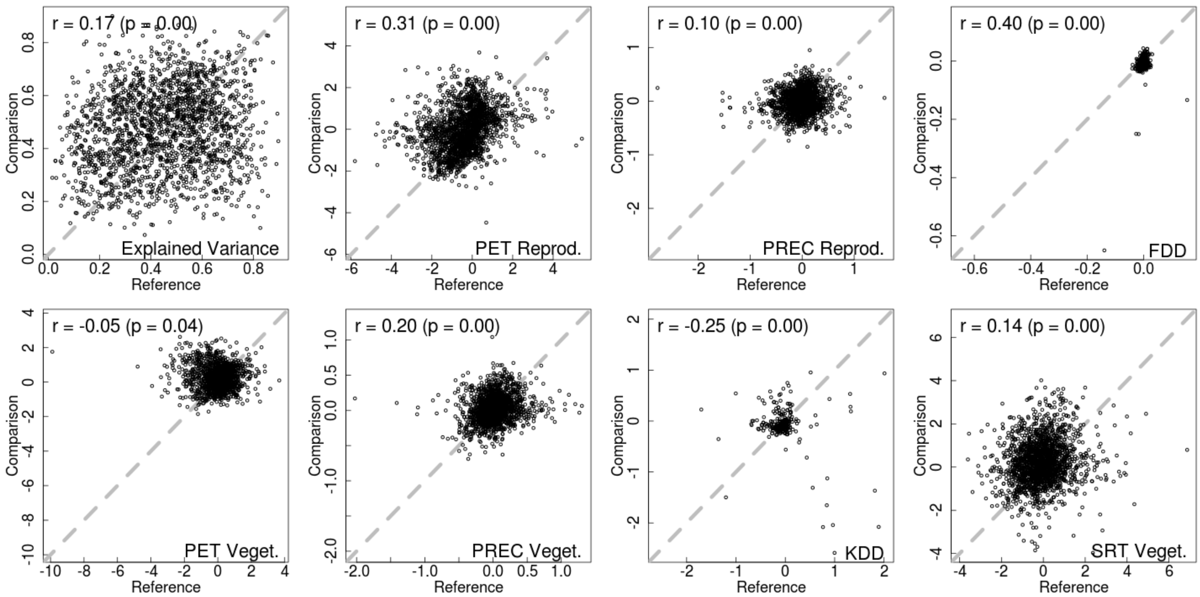


Fig. S7: Comparison of STSM explained variance and coefficients from USDA ("Reference") and GGYD ("Comparison") maize yield data sets. Each point corresponds to the coefficient estimate for one grid cell. Note that all p-values suggest significance, despite the visual impression of practically no correlation. This significance is owed to the rather high number of data points (1,894), which lets even subtle correlations appear significant.

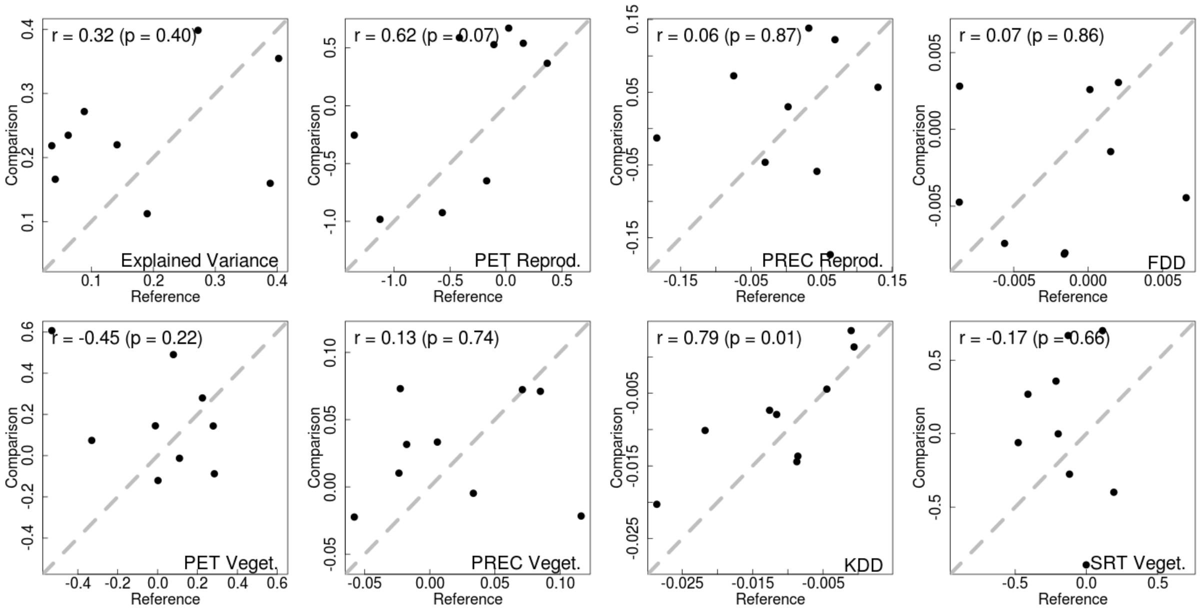
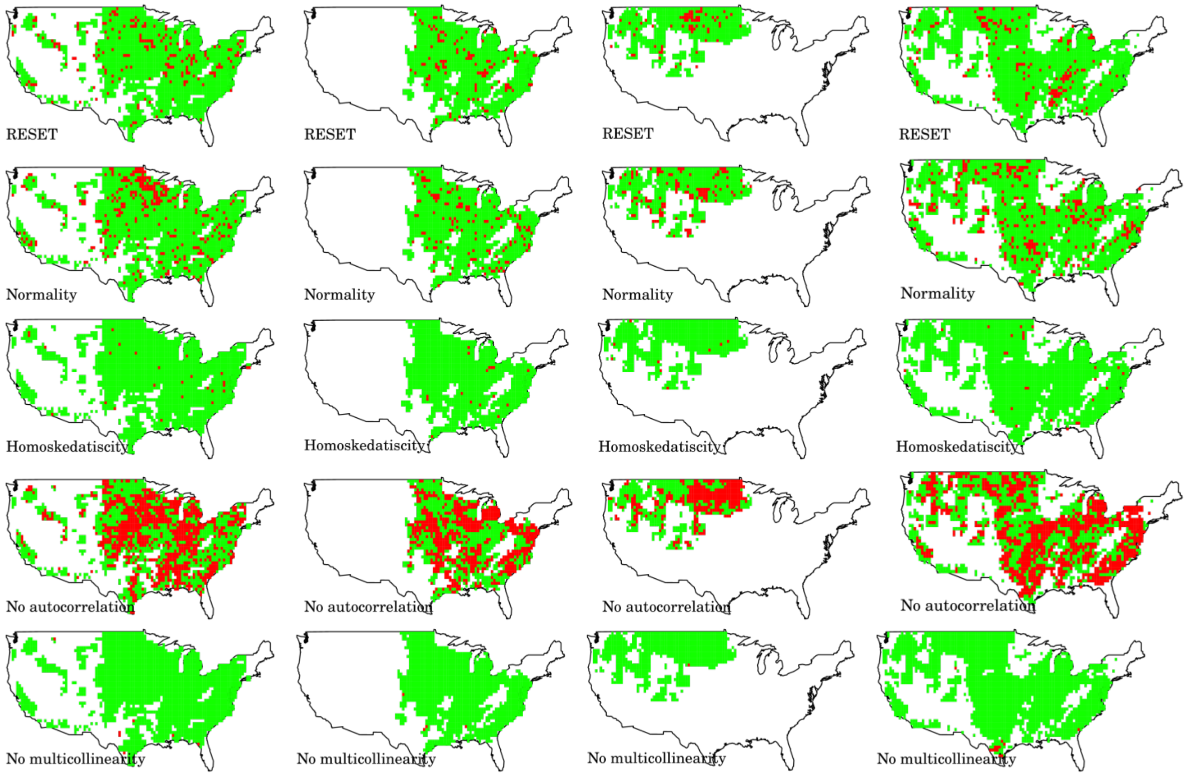


Fig. S8: Comparison of PDM coefficients from USDA ("Reference") and GGYD ("Comparison") maize yield data sets. Each point corresponds to the coefficient estimate for one aggregation region.

11.4.4 Statistical test results

The results of the statistical tests to ensure model validity are displayed in Fig. S9. The RESET test showed that the large majority of grid cells for all four crops were not misspecified, i.e. no quadratic terms were missing. Residuals were normally distributed in most grid cells (Shapiro-Wilk test) and the yield time series were mostly homoscedastic (Breusch-Pagan test). Autocorrelation, however, occurred in a substantial fraction of the grid cells for all crops (Breusch-Godfrey test). This autocorrelation is due to the first difference method and an autocorrelation already in absolute yields (in 55%, 32%, 31% and 37% of grid cells for maize, soybeans, spring and winter wheat, respectively). The LM test for spatial heterogeneity showed, for all crops and all climate regions, that a panel model approach is appropriate (*p* values <0.05; no map provided). The condition index test for multicollinearity following Belsley et al. (1980) showed values above 10 in only 25 out of 5,976 total grid cells (0.4%) for all crops, and all values are below 17. Since only values above 30 would hint to multicollinearity problems we conclude that this is not a problem. Thus, with the exception of autocorrelation no test hints to systematic problems for any of the crops.



(a) Maize **(b)** Soybeans **(c)** Spring wheat **(c)** Winter wheat

Fig. S9: Statistical test results for the US. Columns are crops and rows are the different tests. Green indicates a "successful" test, i.e. no problem, while red indicates a rejection of the respective H_0 of no misspecification/autocorrelation/heteroskedasticity/un-normality. For the condition index test of multicollinearity grid cells are marked in red if there is any variable with a condition index larger than 10.

11.4.5 Combined evaluation of observed yield variability and explained variance

The model explanatory power varies to some extent with the observed yield variability. Yield variability here is measured as coefficient of variation (CV), defined as standard deviation over mean. There are four different combinations: whether the model explains more than 45% of the variation or not, and whether yield variability is substantial (CV >= 0.15) or not. We used the values of 45% and 0.15 to conform with a previous study by Ray et al. (2015). A combined analysis of these four cases is shown in Fig. S10. Regions in green are well explained by the model, with either substantial yield variability (dark green) or not (light green). Regions in blue have low yield variability and this is only less than 45% explained by the model. Regions in red have substantial observed yield variation but the model is not able to capture it. Regions left blank have no harvested area for the respective crop. The fractions of grid cells with substantial variation but low explanatory power of the statistical model (red pixels) are 23%, 23%, 44% and 36% for maize, soybeans, spring wheat and winter wheat, respectively (Tab. S2).

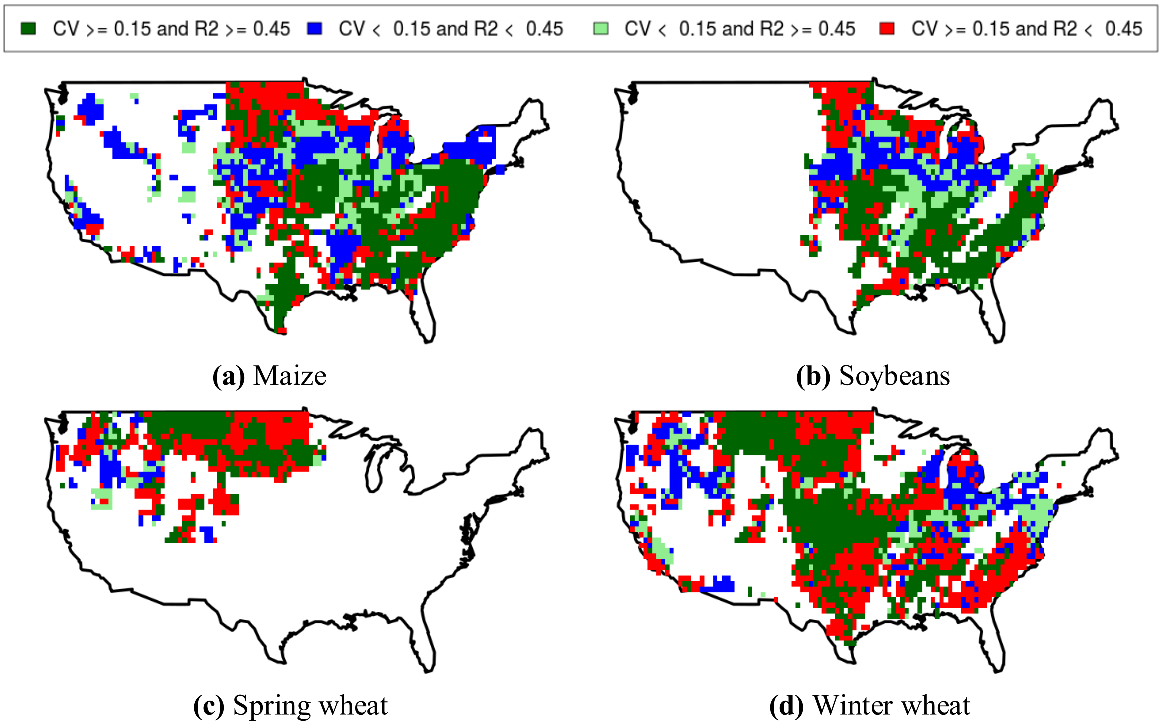


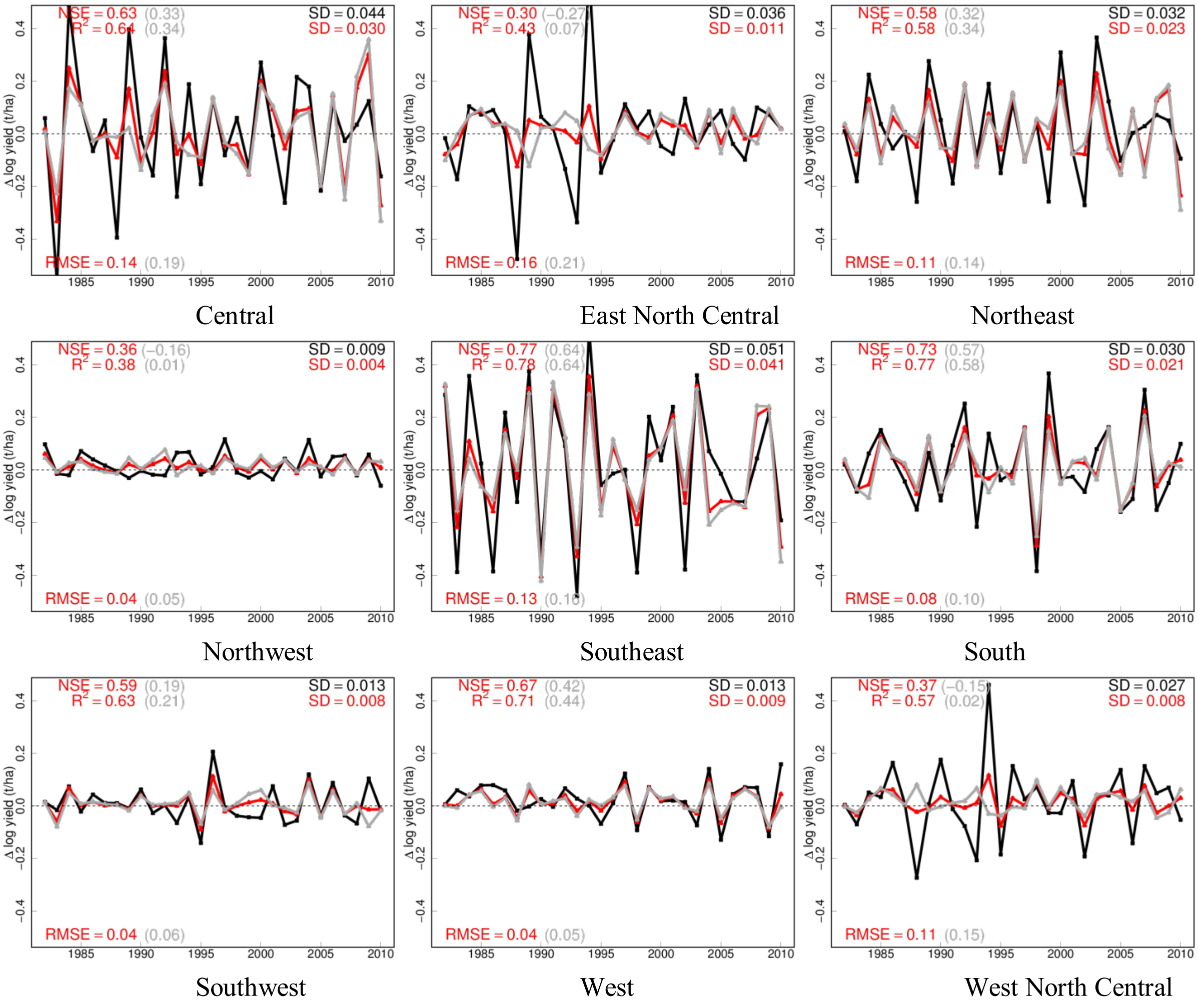
Fig. S10: Combined analysis of model explanatory power vs. yield variation, for USDA maize (panel a), soybeans (b), spring wheat (c) and winter wheat (d).

Tab. S2: Fraction of grid cells (of those where the respective crop is harvested) in different explanation categories (low R^2: < 0.45; low R^2_{O1}: < 0.25; low CV: < 0.15). Numbers in brackets denote analogue fractions for R^2_{O1}. Row sums below or above 100% are due to rounding.

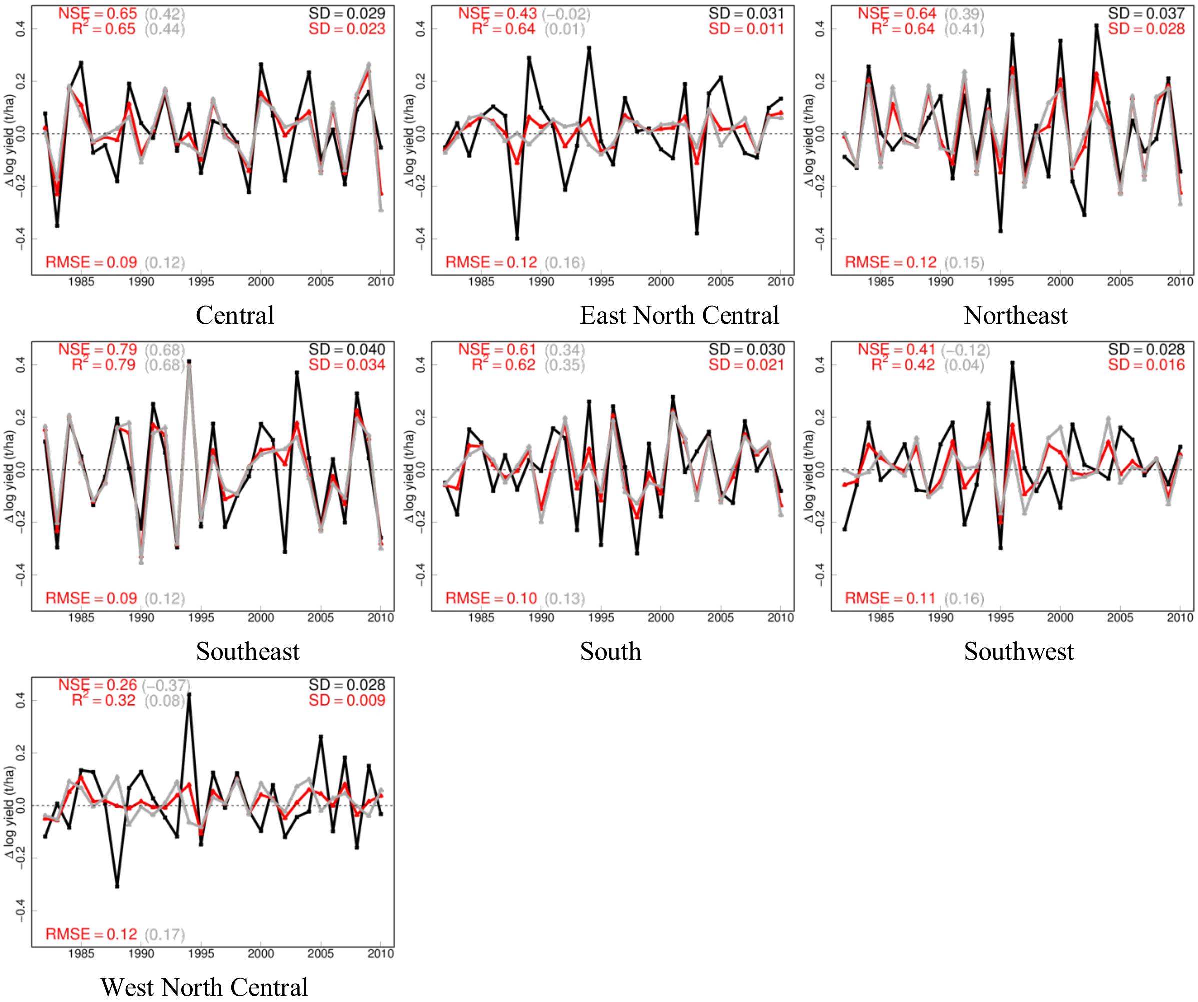
Crop	**Yield data**	**Low CV, low R^2 (R^2_{O1})**	**Low CV, high R^2 (R^2_{O1})**	**High CV, low R^2 (R^2_{O1})**	**High CV, high R^2 (R^2_{O1})**	**Number of grid cells**
Maize	USDA	27 % (33 %)	15 % (9 %)	23 % (35 %)	36 % (23 %)	1,912
Soybeans		20 % (26 %)	17 % (11 %)	23 % (36 %)	40 % (27 %)	1,307
Spring wheat		11 % (14 %)	7 % (4 %)	44 % (64 %)	39 % (19 %)	725
Winter wheat		17 % (20 %)	10 % (6 %)	36 % (50 %)	38 % (24 %)	2,032
Maize	GGYD	40 % (68 %)	48 % (20 %)	4% (8%)	9 % (5 %)	2,021
Soybeans		64 % (84 %)	28 % (9 %)	3 % (6 %)	4 % (1 %)	1,400
Spring wheat		3 % (4 %)	1% (0 %)	42% (65%)	54 % (31 %)	595
Winter wheat		57 % (61 %)	11 % (8 %)	24 % (29 %)	8 % (3 %)	2,036

11.4.6 Time series for US regions

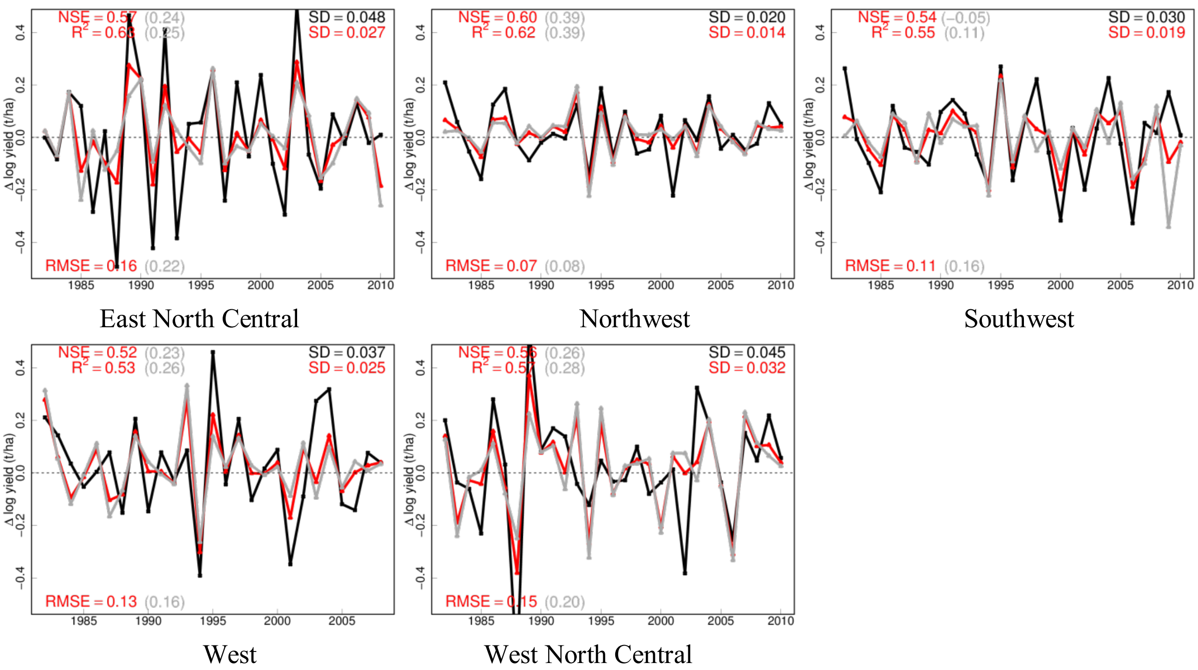
Yield anomaly time series for the nine US climate regions are shown in Fig. S11.



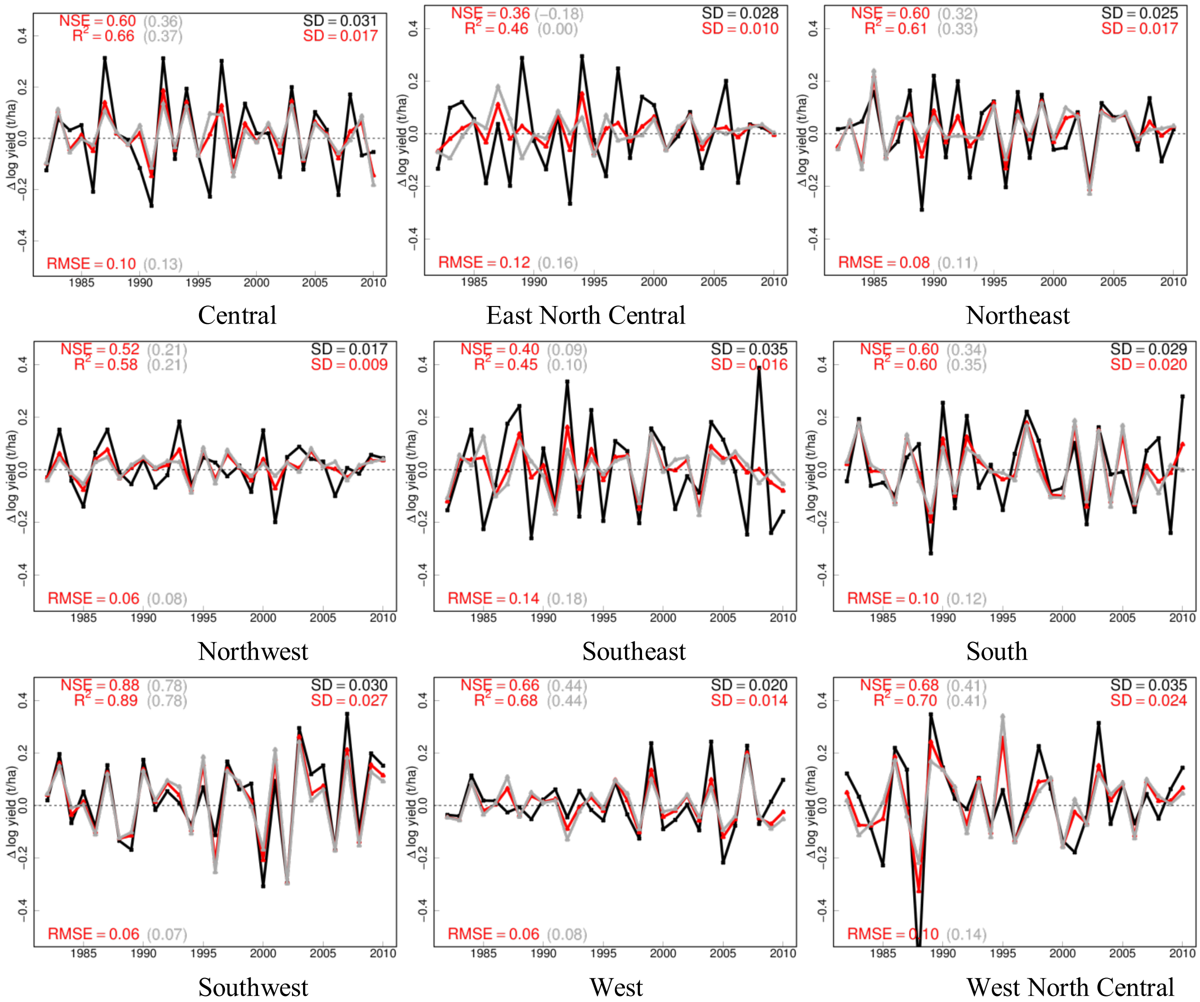
(a) Maize



(b) Soybeans



(c) Spring wheat



(d) Winter wheat

Fig. S11: Yield anomaly time series for the nine US climate regions. Large panels are maize (a), soybeans (b), spring wheat (c) and winter wheat (d). Soybeans and spring wheat are not grown in all regions. Black lines are observed USDA yields, red lines are statistically estimated yields and grey lines are out-of-sample predicted yields. Performance measures printed in red refer to the full model, while grey numbers refer to the out-of-sample estimations.

11.4.7 Model performance for main producers

The list of main producers and the associated three-letter codes are provided in Tab. S3. Countries can be main producers for several crops, leading to 33 unique countries. Correlations between aggregated GGYD and FAO yields are shown in Fig. S12. Yield anomaly time series for the three best reproduced main producers, selected by their R^2_{OI} value, for each crop are shown in Fig. S13. For maize and soybeans the time series for the US (which ranks among the top three) is not shown again (see Fig. 2 of the main paper) such that we resorted to the next ranks. The number of PDM models estimated within each country depends on its size and the availability of GGYD yield data. Subnational zones are defined by administrative boundaries (GADM1; http://gadm.org/). The only exceptions are Russia, which is represented by the three major agricultural areas around the Caspian and Black Sea, and the USA, which is split into the nine climatic zones as defined in Fig. S1.

Tab. S3: Main producers for each crop, sorted by descending total production. Main producers are all countries that together produce more than 90% of world production between 2000 and 2011 according to FAO. The number of subnational regions for PDM estimation, if larger than 1, is indicated in brackets behind the country name.

Maize (24 countries)		Soybeans (5)		Spring wheat (15)		Winter wheat (24)	
Country	**Code**	**Country**	**Code**	**Country**	**Code**	**Country**	**Code**
USA	USA (9)	USA	USA (9)	China	CHN (12)	China	CHN (21)
China	CHN (27)	Brazil	BRA (18)	USA	USA (5)	India	IND (25)
Brazil	BRA (18)	Argentina	ARG (19)	France	FRA (2)	USA	USA (9)
Mexico	MEX (30)	China	CHN (27)	Canada	CAN (5)	France	FRA (21)
Argentina	ARG (19)	India	IND (20)	Australia	AUS (5)	Canada	CAN (7)
India	IND (20)			Turkey	TUR	Germany	DEU (13)
France	FRA (22)			Iran	IRN	Pakistan	PAK
Indonesia	IDN			Poland	POL	Turkey	TUR
South Africa	ZAF			Italy	ITA (19)	Great Britain	GBR
Italy	ITA (19)			Romania	ROU	Argentina	ARG (20)
Canada	CAN (5)			Hungary	HUN	Iran	IRN
Romania	ROU			Syria	SYR	Poland	POL
Hungary	HUN			Russia	RUS (2)	Egypt	EGY
Egypt	EGY			Ukraine	UKR	Italy	ITA (19)
Nigeria	NGA			Kazakhstan	KAZ	Spain	ESP
Philippines	PHL					Romania	ROU
Thailand	THA					Denmark	DNK
Germany	DEU (3)					Brazil	BRA (10)
Spain	ESP					Hungary	HUN
Tanzania	TZA					Syria	SYR
Vietnam	VNM					Morocco	MAR
Ukraine	UKR					Russia	RUS (3)
Russia	RUS (3)					Ukraine	UKR
Kazakhstan	KAZ					Kazakhstan	KAZ

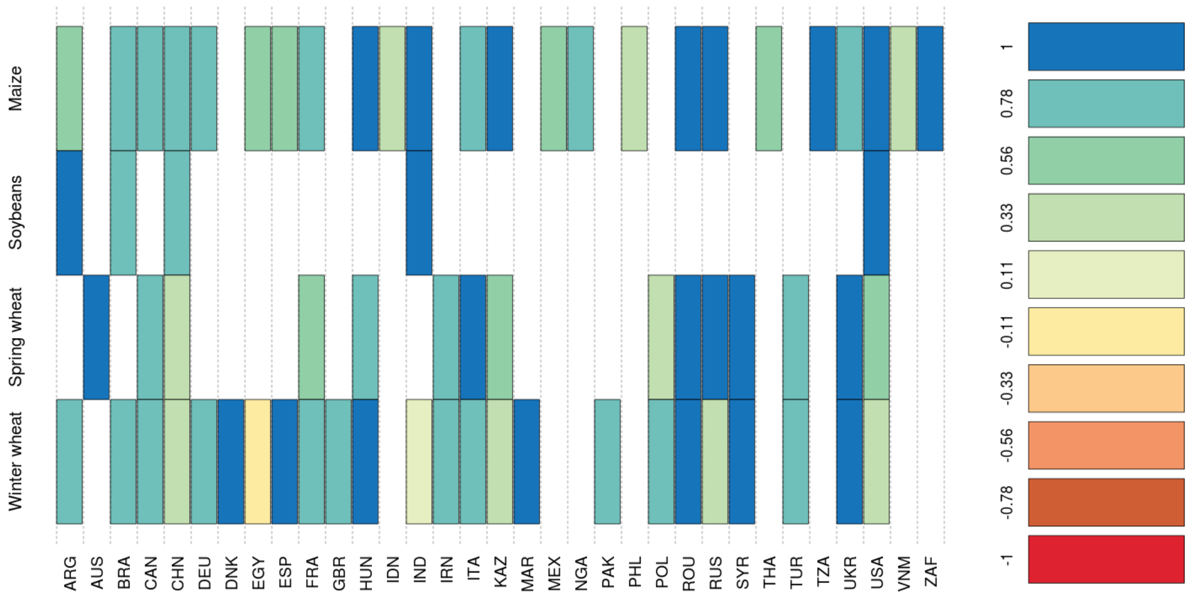


Fig. S12: Correlation (Pearson's r) between nationally aggregated GGYD and FAO national yield anomalies for main producers considered in this study. Yield anomalies were calculated as first differences for both data sets. The MIRCA2000 (Portmann et al., 2010) land-use weighting was applied for aggregation. Applying the M3-Crops harvested areas (Monfreda et al., 2008), which were used for the GGYD construction (Iizumi et al., 2013), instead of MIRCA2000 leads to the same results.

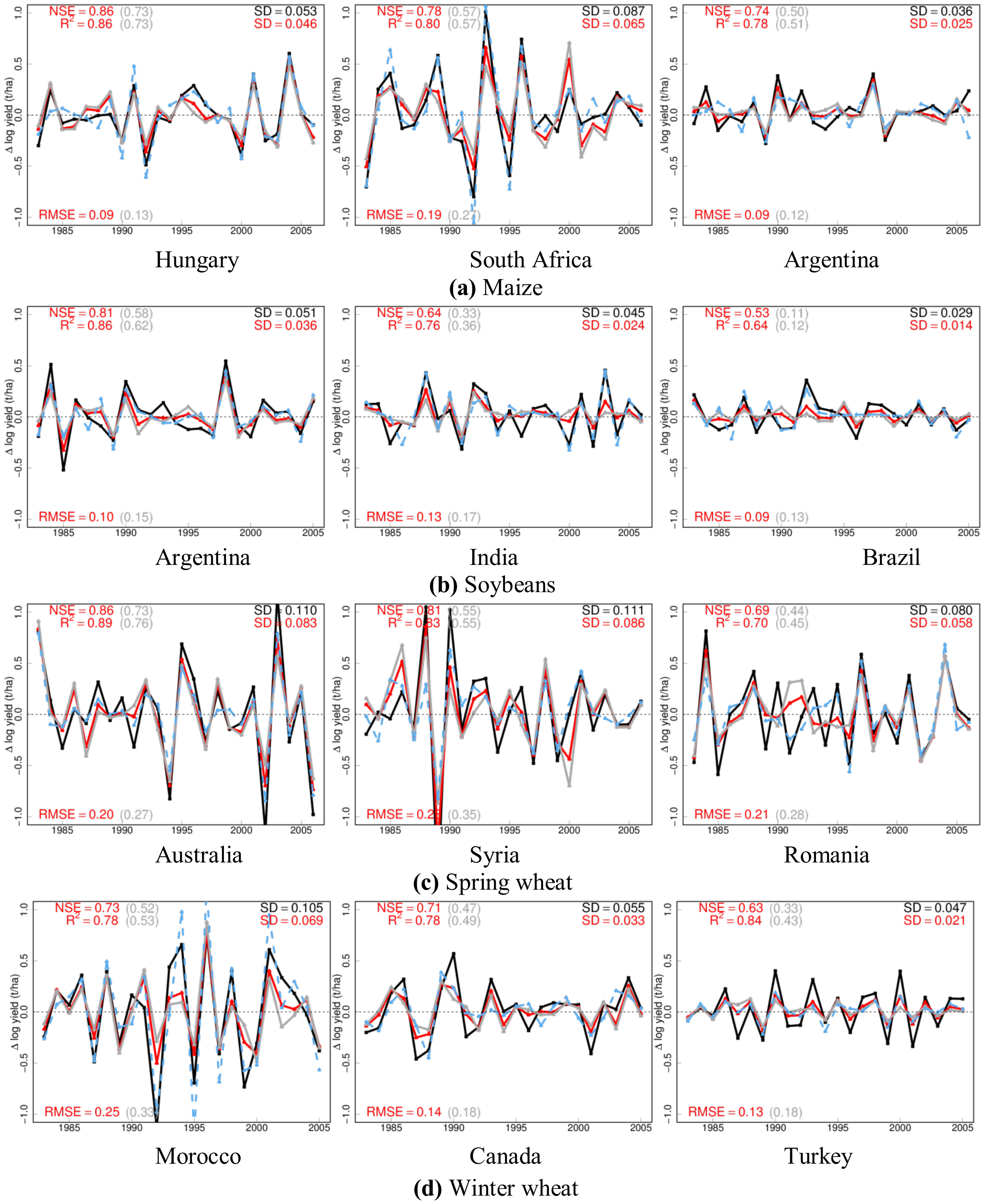
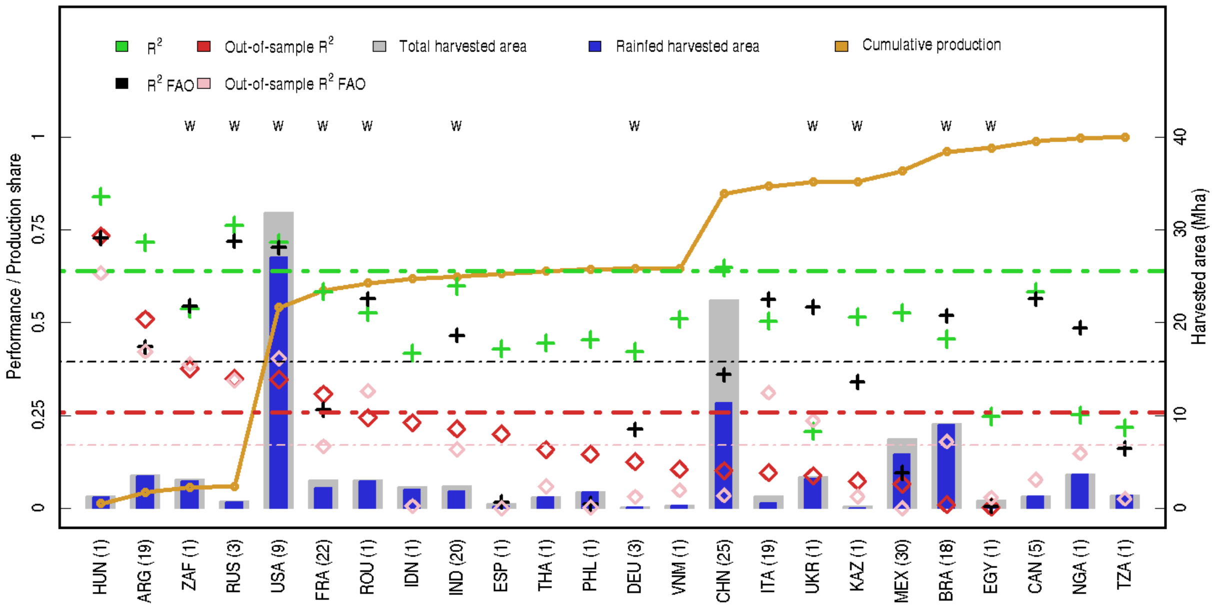


Fig. S13: Yield anomaly time series of selected main producers for maize (a), soybeans (b), spring wheat (c) and winter wheat (d). The three best-performing countries are shown for each crop (excluding the USA).

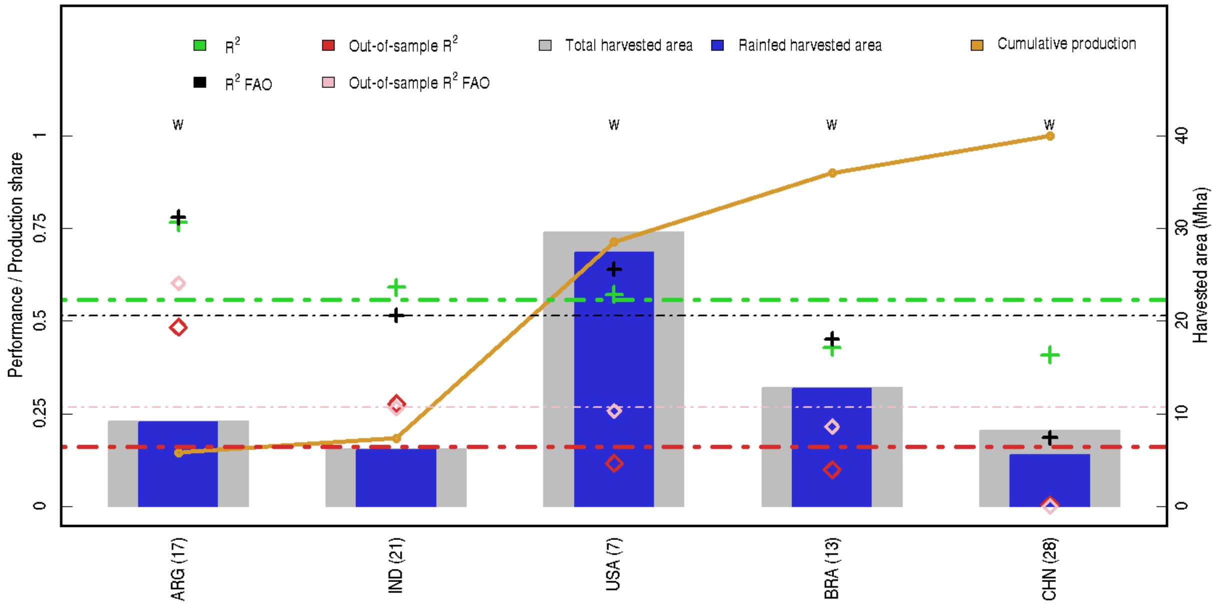
11.4.8 Results for main producers with PDM estimation

Country performances for PDM estimation with GGYD yield data are shown in Fig. S14. The cumulative production share of countries with an R^2_{O1} of at least 25% is 59%, 18%, 46% and 1% for maize, soybeans, spring and winter wheat, respectively. Using the M3-Crops land-use data instead of MIRCA2000 does not change results (data not shown).

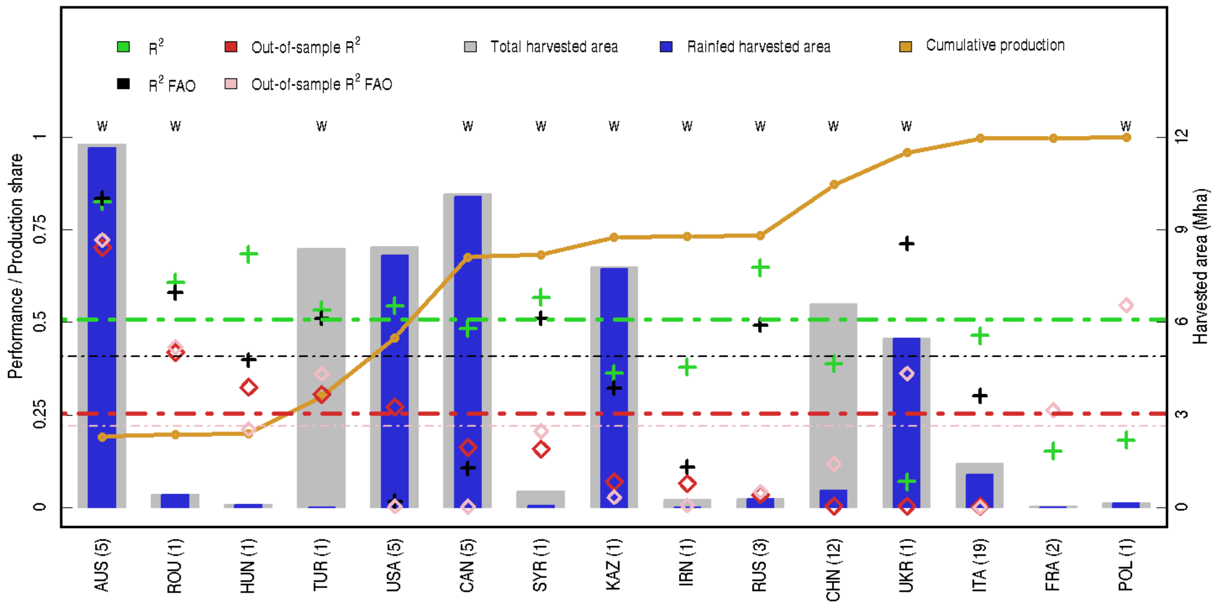
Performance measures differ between FAO and aggregated GGYD yield anomalies (black and green crosses for R^2, or orange and red diamonds for R^2_{O1}). This is expectable since FAO yields were not used for model building and therefore represent a cross-prediction evaluation. The average absolute differences are 19, 8, 21, 23 percentage points for maize, soybeans, spring and winter wheat R^2 values, respectively. For R^2_{O1} these differences are 10, 8, 10 and 7 percentage points.



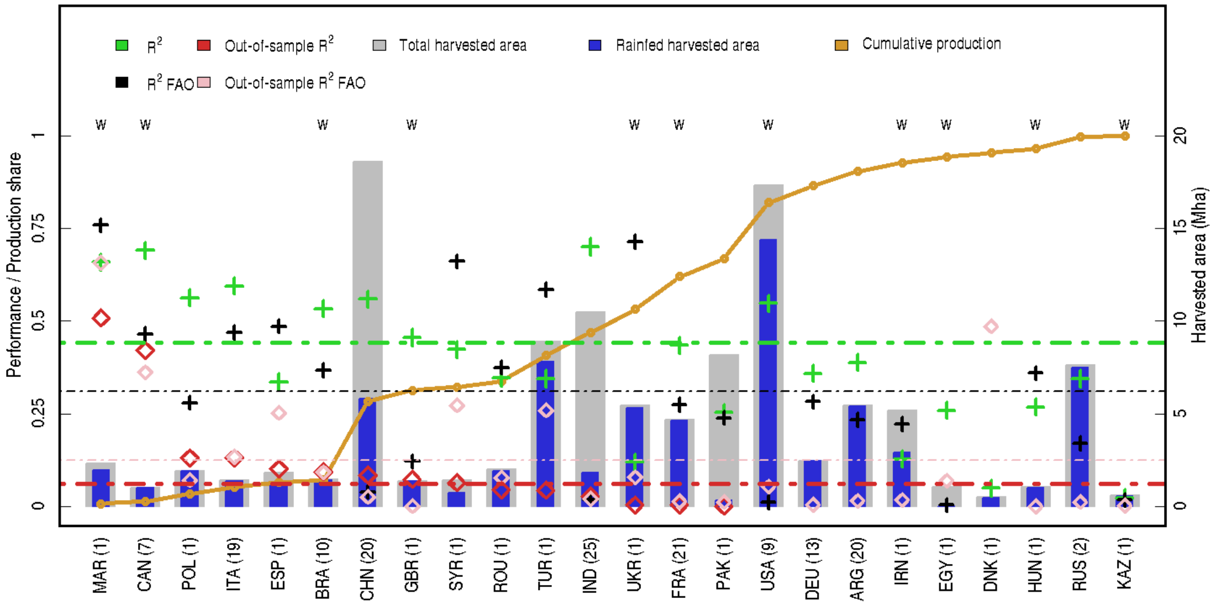
(a) Maize



(b) Soybeans



(c) Spring wheat



(d) Winter wheat

Fig. S14: Model performances in main producers with PDM estimation. Panels are maize (a), soybeans (b), spring wheat (c) and winter wheat (d). Country order is according to descending R^2_{O1}. Colors and lines are as in Fig. 6 of the main paper, with two additional entries: black crosses indicate the R^2 between modeled and FAO yield anomalies, and pink diamonds indicate the FAO-R^2_{O1}, i.e. between out-of-sample predicted and FAO yield anomalies. The number of PDM models estimated per country is indicated in brackets after the country name.

11.4.9 Model performance differences between official yield statistics and GGYD data

The quality of the yield data used for training and evaluating the model is decisive. When using reported yield statistics the out-of-sample performance increases for several countries which otherwise achieve only low performance with the GGYD yield data set (Tab. S4). In some cases, R^2_{O1} is larger for GGYD than the official data set. This happens only, but not necessarily, if the R^2 from GGYD yields is high (>0.85) and higher than the R^2 from official survey data. For soybeans in Brazil the model trained on GGYD yields shows a better performance than with official yield statistics. Possible reasons are a low matching quality between grid cells and Brazilian provinces (which tend to be smaller than one grid cell), or a general inaptitude of the model for (Brazilian) soybean conditions. This will have to be investigated further. Note that Burkina Faso is not a main producer, but subnational crop yield data were available to the authors. All comparisons are based on unweighted aggregation.

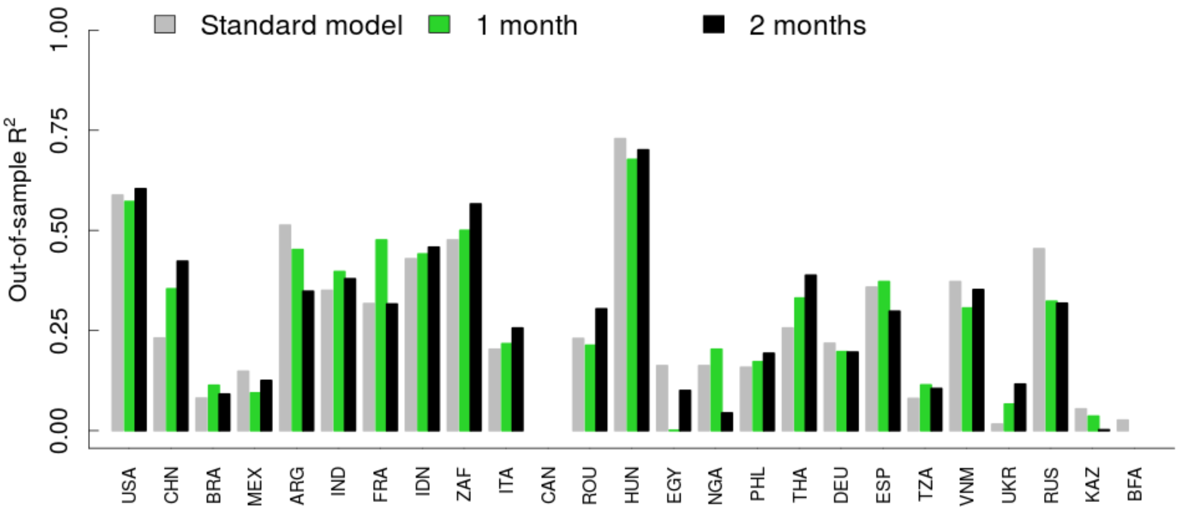
Tab. S4: Different performance of the model when using yield data from statistical offices ("Official") rather than the GGYD data set. R^2_{O1} values which increase by more than 0.1 with official yield statistics are marked in bold.

Country	Crop	R^2 GGYD	R^2 Official	R^2_{O1} GGYD	R^2_{O1} Official	Data source
USA	Maize	0.92	0.81	0.59	0.55	USDA
	Soybeans	0.77	0.69	0.10	**0.45**	
	Spring wheat	0.73	0.63	0.32	0.34	
	Winter wheat	0.92	0.62	0.04	**0.28**	
Germany	Maize	0.62	0.69	0.22	**0.35**[a]	German statistical offices
	Winter wheat	0.44	0.66	n.a. (r < 0)	**0.20**[a]	
Russia	Maize	0.87	0.84	0.45	0.14	Russian statistical office
	Spring wheat	0.67	0.86	0.01	**0.49**	
	Winter wheat	0.59	0.88	n.a. (r < 0)	**0.34**	
Tanzania	Maize	0.68	0.78	0.08	0.16	Tanzanian statistical office
Australia	Spring wheat	0.89	0.85	0.74	0.66	Australian statistical office
Brazil	Maize	0.83	0.89	0.08	**0.73**	Brazilian statistical office
	Soybeans	0.64	0.41	0.12	n.a. (r < 0)	
	Winter wheat	0.71	0.76	0.13	0.15	
Burkina Faso	Maize	0.59	0.71	0.03	**0.43**	Burkina Faso statistical office

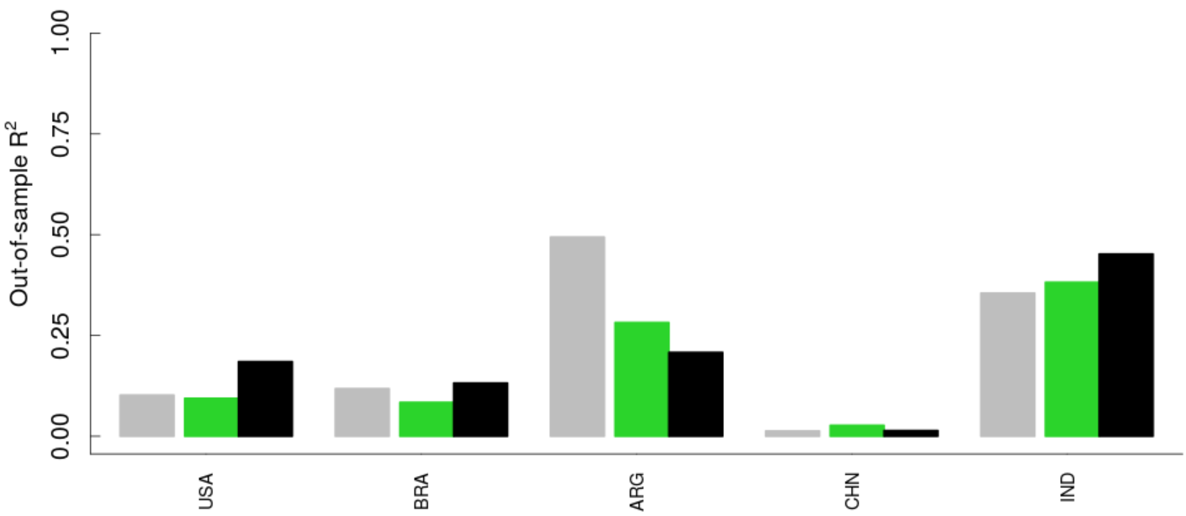
[a] *Note* that one-out-of-sample performances for Germany are higher (0.50 for silage maize and 0.61 for winter wheat) in Gornott and Wechsung (2016) where the model is slightly different and uses different weather data.

11.4.10 Forecasting capacity of the model for all main producers

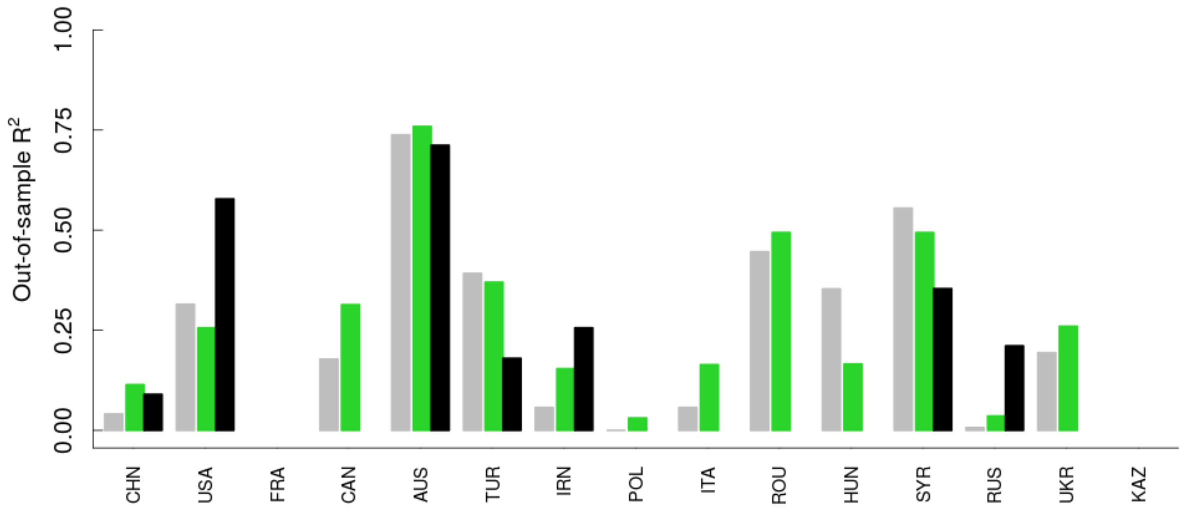
The forecasting capacity of the model, measured by one-out-of-sample R^2_{O1} for prediction with a reduced growing season, is shown in Fig. S15. The share of cumulative production within the main producers that can be predicted with at least 25% accuracy one month before harvest is 82%, 18%, 77% and 11% for maize, soybeans, spring and winter wheat, respectively, and with 50% these are 51%, 0%, 19% and 1%. Two months before harvest the production shares with prediction capacity above 25% are 86%, 4%, 36% and 18% for maize, soybeans, spring and winter wheat, respectively, and with 50% these are 51%, 0%, 35% and 1%.



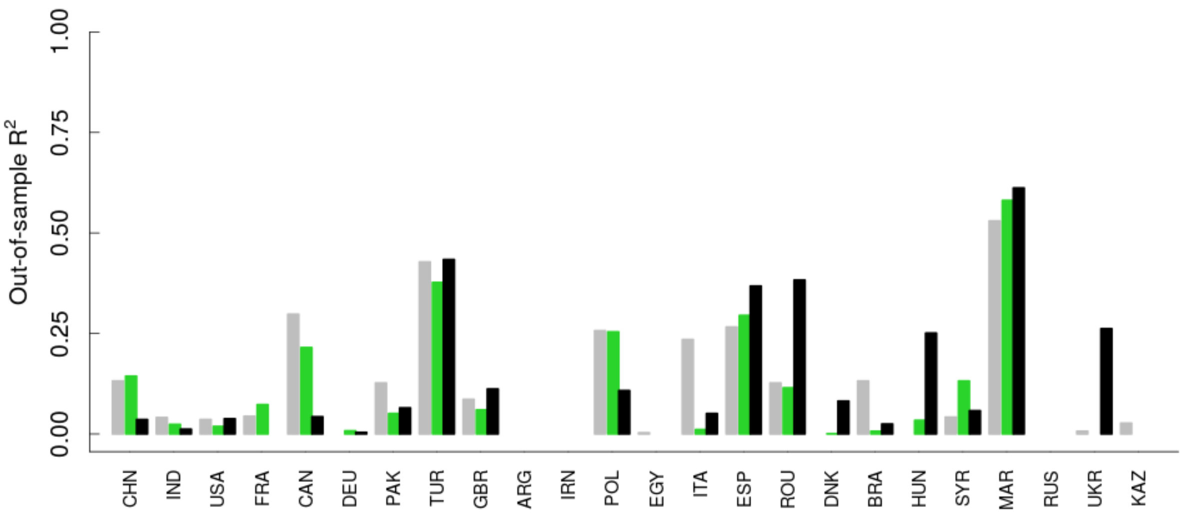
(a) Maize



(b) Soybeans



(c) Spring wheat



(d) Winter wheat

Fig. S15: Forecasting capacity of the model for all main producers, with GGYD yields. Countries are ordered according to descending total production. Gray bars are the standard model with full growing season used for training and prediction. Green and black bars show performance when withholding one or two months, respectively, for training the model and predicting yield anomalies out of sample. Note that in some cases no performance data are present, for either of two reasons: the reduction of the growing season did not allow for calculating any regression, or the correlation between observed and predicted anomalies is negative and thus its squared value would be misleading.

11.4.11 References

Belsley DA, Kuh E, Welsch RE (1980) *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity,* New York, Wiley.

Gornott C, Wechsung F (2016) Statistical regression models for assessing climate impacts on crop yields: A validation study for winter wheat and silage maize in Germany. Agricultural and Forest Meteorology*,* 217, 89-100.

Iizumi T, Yokozawa M, Sakurai G *et al.* (2013) Historical changes in global yields: major cereal and legume crops from 1982 to 2006. Global Ecology and Biogeography*,* 23, 346-357.

Karl TR, Koss WJ (1984) Regional and National Monthly, Seasonal, and Annual Temperature Weighted by Area, 1895-1983. In: *Historical Climatology Series 4-3*, Asheville, NC, National Climatic Data Center.

Monfreda C, Ramankutty N, Foley JA (2008) Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. Global Biogeochemical Cycles*,* 22, n/a-n/a.

Portmann FT, Siebert S, Döll P (2010) MIRCA2000-Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. Global Biogeochemical Cycles*,* 24, GB1011.

Ray DK, Gerber JS, Macdonald GK, West PC (2015) Climate variation explains a third of global crop yield variability. Nat Commun*,* 6, 5989.

11.5 Covering smallholder farmers' weather perils – a crop model based insurance approach for Tanzania

Christoph Gornott[1*], Fred Hattermann[1], and Frank Wechsung[1]

[1] Potsdam Institute for Climate Impact Research (PIK)

[*] Corresponding author

11.5.1 Materials and methods

11.5.1.1 Data

11.5.1.1.1 Yield information

We use observed farm maize yields supplied by the Ministry of Agriculture, Food Security and Cooperatives (MAFSC, 2010) for Tanzanian districts ($N = 116$) and the period 2003 to 2010. These observed yield data contain some implausible outliers, which are far beyond the genetic yield potential of maize. Schlenker and Lobell (2010) show that the direct use of such observed yields can amplify the uncertainty of yield assessments. Thus, we eliminate implausible outliers in the observed yield dataset by using reasonable upper yield ceilings. The ceiling is 25% above the local yield calculated by our process-based model assuming a fertilization of 120 kg N and 40 kg P_2O_5 ha^{-1}. After adjusting the yield dataset (removing implausible outliers or too short time series), we still work with N = 104 districts. Since this dataset has also some missing values, in total our dataset contains 796 observed yield values. The average of the original dataset is 1.5 t ha^{-1} and the standard deviation is ±2.0 t ha^{-1}, while the average of the adjusted dataset is 1.3 t ha^{-1} with a standard deviation of ±0.9 t ha^{-1}.

11.5.1.1.2 Weather information

We use reanalyzed weather information (WFDEI ERA-Interim) of 319 grid points across Tanzania from 1979 to 2012 (Weedon et al., 2014). To justify the usability of the dataset, we compare the reanalyzed weather dataset with nearest weather data from 16 stations (Tanzania Meteorological Agency, 2007) of the period 1970 to 2006. The plots in Fig. S.1 exemplarily show the comparison of the yearly-averaged, intra-annual precipitation distribution for five reanalyzes points and observed weather stations. In general, the seasonality of the observed weather data is reproduced by reanalyzed weather data for all weather stations. The 6-day moving average shows that the reanalyzed weather data satisfactorily represent the measured data: The Nash–Sutcliffe model efficiency coefficient (NSE, see also Chipanshi et al. (2015) for the calculation) for precipitation (*PREC*) is: 0.77, NSE for minimum temperature (TMP_{min}): 0.74, and NSE for maximum temperature (TMP_{max}): 0.31.

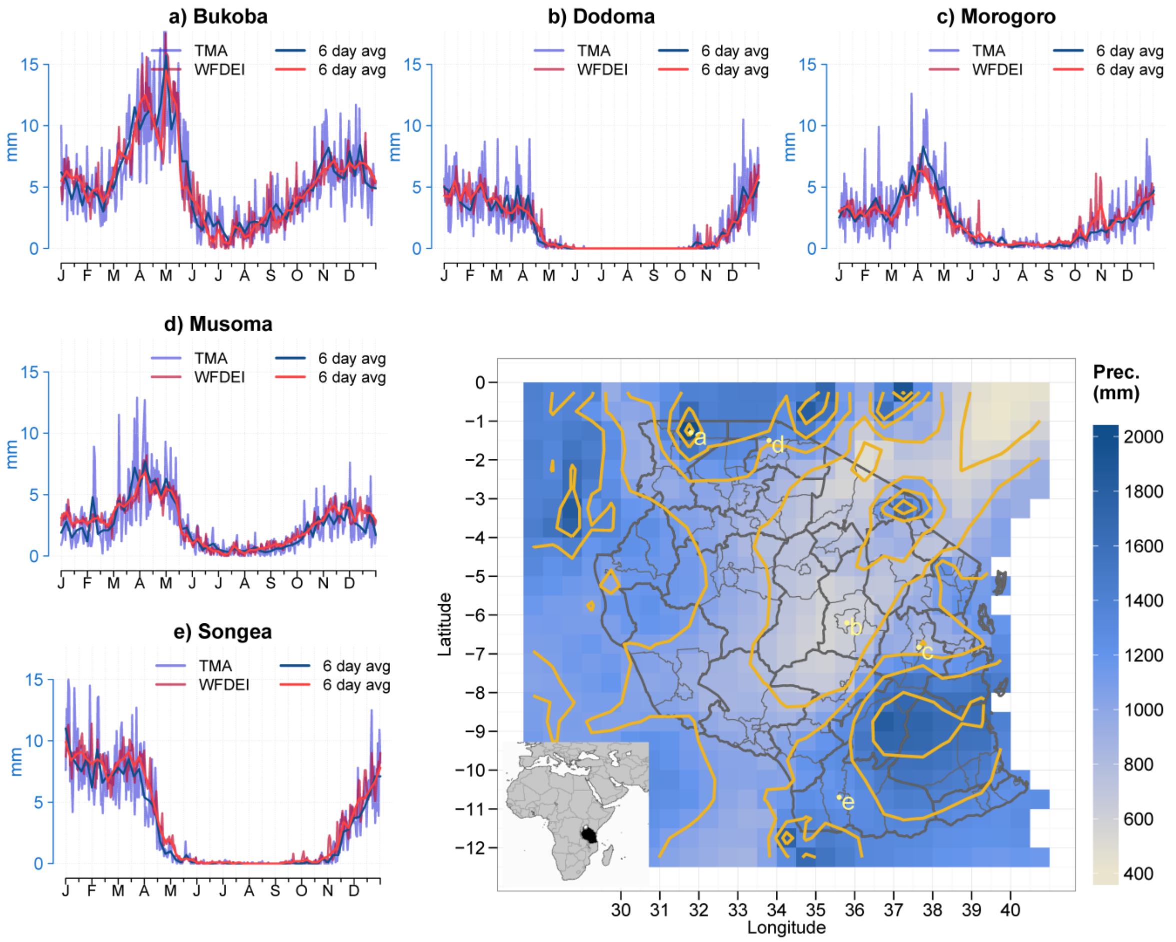


Fig. S.1. Spatial and intra-annual distribution of reanalyzed and observed precipitation patterns. The iso-precipitation lines are yellow; the black boundaries are the regions (thick lines) of Tanzania with its districts (thin lines). The weather stations are the white points a–e. The acronym WFDEI stands for the reanalyzed and TMA for the observed weather data.

11.5.1.1.3 Soil and agronomic management information

The soil information for the 319 weather grid points is taken from the FAO-74 soil classification according to Dewitte et al. (2013) and the ILRI (2005) soil map. We use the fertilization amounts according to Thornton et al. (2009) as input for the process-based model. For the statistical model, we use the variables acreage maize (district scale), paid agricultural subsidies, and urea application (both on national scale) provided by the Ministry of Agriculture, Food Security and Cooperatives (MAFSC, 2010). Finally, we use agro-ecological zones (IFPRI, 2015) for the classification of semi-arid and sub-humid regions (Fig. S.2). The division of the maize growing season (planting to harvesting periods) is taken from FAO Crop Calendar (2010). For maize prices, we use the national price statistics from the FAO Stat (2013).

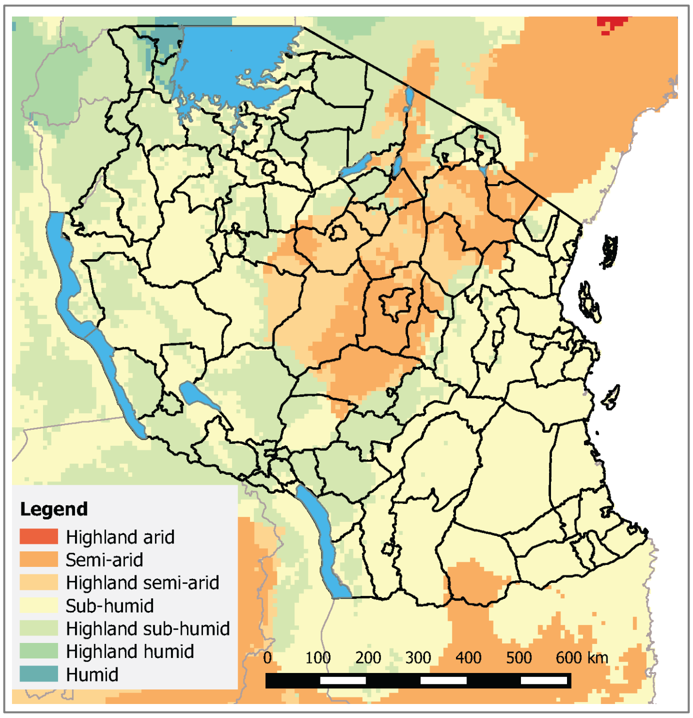


Fig. S.2. Agro-ecological zones and Tanzanian districts.

11.5.1.2 Process-based crop modeling

We use the Soil and Water Integrated Model (SWIM) as process-based model. The processes of SWIM are calculated on a daily time step for spatial points, which are representative for larger regions (sub-national boundaries, hydrotopes, field trials, or grid cells) (Krysanova et al., 2015, 2000). For our investigation, we apply SWIM on grid cell information of 0.5° (approximately 50km at the equator).

11.5.1.2.1 Crop yield modeling by SWIM is based on the EPIC crop module

EPIC is a worldwide applied process-based crop model (Asseng et al., 2013; Rosenzweig et al., 2014), which is able to reproduce the cropping conditions in SSA (Folberth et al., 2012). The model computes crop yields as a product of the total above-ground biomass and the harvest index. While the harvest index increases until harvesting, the above-ground biomass growth is calculated as the product of the crop-specific parameter for converting energy to biomass and the photosynthetic active radiation. The photosynthetic active radiation is a function of the incoming solar radiation and the leaf area index (LAI) of the corresponding crop. Any divergence of these optimal growing conditions reduces the biomass growth by the stress factors heat stress and insufficient water, nitrogen, and phosphorus supply within a minimum function. The plant water supply is determined by precipitation and water withdrawn by evaporation, surface runoff, infiltration, and plant water uptake, respectively. The EPIC crop module embeds a nitrogen and phosphorus cycle. The nitrogen cycle includes mineralization, nitrification, and denitrification. The phosphorus cycle includes phosphorus adsorption and mineralization. The nitrogen and phosphorous supply is added by organic and mineral fertilization.

11.5.1.2.2 Modifications in the EPIC crop module

In the EPIC crop module within SWIM, we use mostly the standard maize parametrization of the EPIC model (Krysanova et al., 2000). For Tanzanian maize yield assessments, we adjust the temperature sensitivity, harvest index, maximum LAI, and required heat units to maturity. Folberth et al. (2016, 2013, 2012) show that this parametrization is valid for the cropping conditions in SSA. The crops in SWIM are not parametrized on one or more individual crop varieties. Information about crop varieties would be very helpful for an accurate representation of local or regional cropping conditions. However this information seems to be unavailable for Tanzania. The temperature sensitivity is corrected according to Rötter and van de Geijn (1999) to 8 °C basic and 28 °C optimum temperature. According to Gaiser et al. (2010) and McClung (McClung, 2014), the harvest index of local, unimproved crop varieties is lower than for improved varieties. Since seed saving of local varieties is common in Tanzania (Nkonya and Mwangi, 2004; Westengen et al., 2014), we take a HI of 0.35 (Gaiser et al., 2010). Folberth et al. (2013) show that a HI parameter of 0.35 leads to reasonable results for entire SSA. Depending on the environmental and crop genetic conditions, the maximum LAI varies highly in SSA (Lukeba et al., 2013). Following Gaiser et al. (2010), we use a maximum LAI of 6.0 m^2 (leaf) m^{-2} (ground). The maize maturity groups are covered by the heat units. The heat units are the accumulated growing temperature (actual temperature reduced by the basis temperature) sum from seeding to maturity of the crop. For Tanzania as a whole, we use medium-maturity varieties with 2800 °C heat units (FAO, 2015; Folberth et al., 2012). In our model, the management is uniform for all grid points. The variables harvesting dates, nitrogen dynamics in the soil (e.g., leaching), or other soil properties (e.g., water holding capacity, rooting depth) vary across space (grid-specific) in our process-based model. Planting dates and fertilizer application are uniform across districts. The planting date is set relatively early (December 10th). In the process-based model, the plant germination starts with the first rains and the plant will not die within the first 30 days also with insufficient water supply. Due to this, we implicitly account for differences in the planting dates. The harvest date is 6 days after maturity (or after this, the next day without precipitation). According to the World Bank (2016) survey, the average Tanzanian maize fertilization is 23.0 kg N and 0.0 kg P_2O_5 ha^{-1}. For smallholder farmers, Thornton et al. (2009, p. 57) describe an inorganic fertilization of 5.0kg N and no phosphorus fertilization. Following the later, we apply an inorganic fertilization of 5.0 kg N and 0.0 kg P_2O_5 ha^{-1}, because the fertilization is rather poor than sufficient (Tittonell and Giller, 2013; Vitousek et al., 2009). In particular in semi-arid, tropical regions, the nutrient uptake is highly influenced by the soil moisture. According to Folberth et al. (2012) and Harmsen (2000), we included in the Liebig minimum function an interaction between water and nutrient stress to calculate the crop growth regulating factor (*GRF*, Eq. S.1). The fertilization is applied 13 days after sowing.

$$GRF = \min(TMP^{Stress}, Water^{Stress}) \quad \min\left(N^{Stress}, P_2O_5^{Stress}\right) \tag{S.1}$$

11.5.1.3 Statistical crop models

We use an approach, which is similar to the statistical regression model introduced by Gornott and Wechsung (Gornott and Wechsung, 2016) for the case of Germany. For our approach, we use the same conceptual framework for the variable selection and the same statistical methods and consider both non-weather and weather impacts on maize yields. However, we use a different functional form and variable transformation, which fits better to the Tanzanian weather and agronomic conditions.

11.5.1.3.1 Statistical method

We use three different statistical regression methods to capture the spatial and temporal heterogeneity among *N* districts and *T* years: separately estimated time-series models (STSMs), panel data models (PDMs), and random coefficient models (RCMs). The STSMs estimate independently a separate time-series model for all districts. Each STSM explains the yield variability by a district-individual intercept and district-individual parameters. The PDMs capture directly temporal and spatial variability by one parameter set valid for *N* considered districts. RCMs contain both one parameter set for all *N* districts and district individual parameters. Since these parameters depend on each other, the RCMs are estimated by the restricted maximum likelihood method instead of the ordinary least squares method used for STSMs and PDMs.

11.5.1.3.2 Modeling approach

Combined process-based and statistical modeling approach to assess weather and non-weather-related yield variability by using a logarithmic function ($\mathbf{PM^{we}}$-$\mathbf{SM^{nw}}$-$\mathbf{SM^{we2}}$): The process-based model (PM) is supposed to explain the direct weather-related yield variability (we). The residual yield variability ($\varepsilon_{it} = y_{it} - y_{it}^{PBM}$) of the observed ($y_{it}$) and process-based modeled yields (y_{it}^{PM}) is explained by (i) non-weather-related (nw) and (ii) indirect (second-order) weather-triggered yield influences (we2) like pests and diseases. (iii) Since our dataset only has limited degrees of freedom ($T = 8$), we estimate non-weather-related and indirect weather-triggered impacts in two consecutive statistical models. This means that the non-weather-related statistical model uses the residuals of the process-based model. In the consecutive step, the weather-driven statistical model uses the residuals of the non-weather-related statistical model as the endogenous variable. This approach enables the consideration of both impact factor groups without the risk that any impact factor is considered twice. While the STSMs are directly estimated on district scale, the PDMs are estimated on regional scale. Due to this, PDMs have more available degrees of freedom. (iv) This allows the consideration of both non-weather-related and indirect weather-triggered impacts in a single PDM. RCMs require the same amount of degrees of freedom as the STSMs, thus, they are applied like the STSM for approach i–iii. In total, we compare four approaches to explain the PM residuals (ε^{we}): (i) only i non-weather-driven, (ii) only weather-driven, (iii) non-weather and weather-driven (in two consecutive statistical models), and (iv) non-weather and weather-driven in one model:

(i) PM^{we}-SM^{nw} (ε^{we}): one statistical model to capture the non-weather-related impacts on ε^{we} (Fig. S.3),
(ii) PM^{we}-SM^{we2} (ε^{we}): one statistical model to capture indirect weather-triggered impacts on ε^{we},
(iii) PM^{we}-SM^{nw} (ε^{we})-SM^{we2} (ε^{nw}): two consecutive statistical models to captures non-weather-related impacts on ε^{we} and the residuals of that model (ε^{nw}) by indirect weather-triggered impacts (Fig. S.4),
(iv) PM^{we}-SM^{nw}&SM^{we2} (ε^{we}): one statistical model to capture both non-weather-related and indirect weather-triggered impacts on ε^{we} (only investigated with PDMs).

We use a logarithmic function as the basic functional form with the residuals (ε_{it}) as endogenous variables (Eq. S.2). The exogenous variables are either the non-weather-related or indirect weather-triggered variables. The J exogenous variables are transformed to logarithmic values. The terms β are the parameters, u is the error term, t (with $t = 1, \dots, T$) is the time-index, and i denotes the spatial index (with $i = 1, \dots, N$). The endogenous variable is considered in untransformed values, because the negative residual values allow no logarithm. We use exogenous variables as logarithmic values, because this transformation achieves better results than the untransformed ones. We also investigate several other transformations (see S.2.3.3), however, the fixed effects transformation achieves the best goodness of fit.

$$\varepsilon_{it}^{id/we2} = \log \beta_{0i} + \sum_{j=1}^{J} \beta_{ji} \log \, x_{jit} + \log u_{it} \tag{S.2}$$

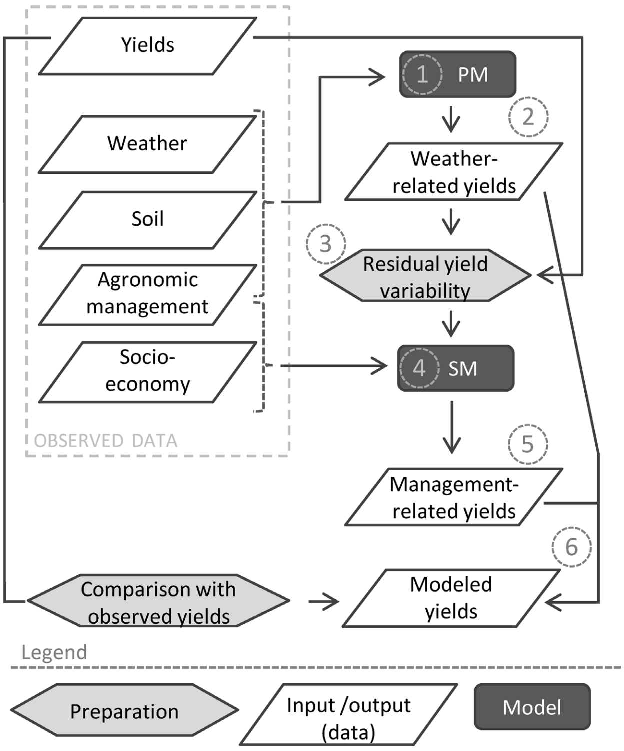


Fig. S.3. Flowchart of the combined model approach (PM^{we}-SM^{nw}): (1) Application of the process-based crop model (PM) for the region-specific agro-ecological conditions. (2) Separation of weather-related yield variability and (3) calculation of residual yield variability. (4) Application of the statistical crop model (SM) to capture the residual, non-weather-related yield variability by agronomic and socio-economic impacts. (5) Separation of non-weather-related yield variability. (6) Combination of weather-related and non-weather-related yield variability to compare the modeled yields with the observed yields.

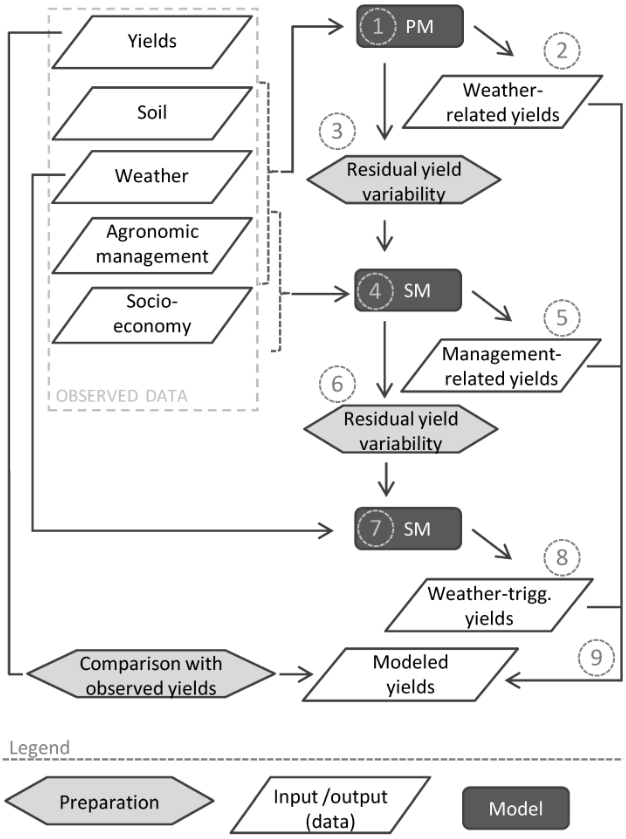


Fig. S.4. Flowchart of the combined model approach (PM^{we}-SM^{nw}-SM^{we2}): (1) Application of the process-based crop model (PM) for the region specific agro-ecological conditions. (2) Separation of weather-related yield variability and (3) calculation of residual yield variability. (4) Application of the statistical crop model (SM) to capture the residual, non-weather-related yield variability by agronomic and socio-economic impacts, (5) separation of non-weather-related yield variability. (6) Calculation of residual yield variability (from the SM). (7) Application of the statistical crop model (SM) to capture the residual, indirect weather-triggered yield variability (5) separation of indirect weather-triggered yield variability. (9) Combination of weather-related, non-weather-related and indirect weather-triggered yield variability to compare the modeled yields with the observed yields.

11.5.1.3.3 Variable selection

In general, we use the same weather and non-weather exogenous variables for our statistical model approaches. The approaches are driven with a set of either non-weather or weather variables. As non-weather exogenous variables, we use *maize acreage* (in ha per district), *urea application* (in metric tons for entire Tanzania), and *paid subsidies on crop production* (in US$). The *maize acreage* is thought to capture agronomic management decisions, land use and land availability (Iizumi and Ramankutty, 2015). This includes crop rotation preferences (for more maize) and the economic suitability of maize production. Moreover, yield can be interpreted as land productivity (Kaufmann and Snell, 1997) and changes in acreage might go ahead with changes in soil quality, because farmers plant on marginal soils in case of an acreage expansion (Cassman et al., 2003). This will have a direct impact on crop yields. *Urea application* (in metric tons for entire Tanzania) should cover fertilizer availability and application (van der Velde et al., 2014). The application of fertilizer is already modeled by the process-based model. However, because information about annual variation of fertilizer application is not available at district level, fertilizer application is kept constant over time and only used to reproduce the average yield level. As a result annual yield variation attributable to fertilizer application cannot be reproduced by the process-based model. Thus, the statistical model is needed to explain also the yield variability attributable to changes in fertilizer application. For SSA, Ward et al. (Ward et al., 2014) use as economic variables nitrogen and phosphorus fertilizer as well as irrigation. Since only 1.8 – 3.3% of the Tanzanian cropland is irrigated (NBS, 2012, p. 107; You et al., 2011), we haven't included this variable. Other production factors, e.g., machinery (You et al., 2009), also seems to be of lower importance for Tanzania. Pauw and Thurlow (Pauw and Thurlow, 2011) show that agricultural growth stagnates since the 1990s, because of low investments in infrastructure and machinery. The variable *paid subsidies on crop production* (in US$) addresses the efficiency of the fertilizer and seeds' subsidy system and the socio-economic behavior of farmers (Jayne et al., 2013). Fertilizer and seed subsidies tripled the maize yields in Malawi within three years (Sánchez, 2010). Such subsidies are also disbursed in Tanzania, however, with a smaller yield increase (Benson et al., 2012). Tanzania has launched an input subsidy program in 2003 with the main objective to facilitate fertilizer and improved seeds' use in remote areas. This program was changed in 2008 with the aim to raise maize and rice production. The program was designed to increase Tanzania's household and national food security and to response to the fertilizer prices spikes in 2007 and 2008. Because of the disbursed subsidies for improved seeds and fertilizer, the farmers have adjusted their agronomic management with direct implications on crop yields (Minot and Benson, 2009).

Since market access, land tenure security or access to extension are rather time-constant in Tanzania (Deininger et al., 2017; Dillon and Barrett, 2017), the impact of these variables is captured by the intercept of our statistical model and the other parameters are not biased (no omitted variable bias). The spatial heterogeneity of these variables is also captured by the district-individual intercepts. Thus, our model accounts for time-invariant and spatial heterogeneous impacts of market access, land tenure security or access to extension.

As weather variables, we use *solar radiation*, *precipitation,* and the *vapor pressure deficit*. The solar radiation (*SR*) maps the potential growth. The variables *PREC* and vapor pressure deficit (*VPD*, Eq. S.3 for calculation) should capture deviations from the optimal water supply. The *VPD* is calculated by $TMP_{\max}$ and $TMP_{\min}$ (Castellvi et al., 1997). As indirect weather-triggered impacts, these variables should address plant health (pests, weeds, and diseases) and agronomic management, which is collinear with the weather variables (Rosenzweig et al., 2001).

$$VPD = 6.11\left(\exp^{\left(\frac{17.269\,TMP_{\max}}{237.3\,+\,TMP_{\max}}\right)} - \exp^{\left(\frac{17.269\,TMP_{\min}}{237.3\,+\,TMP_{\min}}\right)}\right) \tag{S.3}$$

The Tanzanian maize growing season lasts approximately from December to June. Within Tanzania there is a high heterogeneity of the planting and harvesting dates (FAO Crop Calendar, 2010). Therefore, the weather variables are aggregated by the district-specific maize growing season. Because of limited degrees of freedoms, the variables are not further divided in sub-parts of growing season.

11.5.1.4 Validation

To test the robustness of the process-based and statistical models, we conduct a validation with observed yields, which are not considered in the model calibration. Process-based models endogenously compute crop yields (without the consideration of observed yields). This allows for a direct comparison of the observed and modeled yields. Statistical models require observed yields for their estimation. To validate statistical models with unconsidered observed yields, we apply an out-of-sample cross-validation. This validation reduces the estimation dataset by the values of the year t subsequently for all years T. For each year, the parameters are estimated for the reduced dataset (validation parameter). Finally, the yields of the removed years are calculated by the validation parameter and the exogenous variable values of the removed years (Chipanshi et al., 2015).

11.5.1.5 Aggregation of results

The yields of the process-based model are calculated on grid scale and aggregated from grid-cell scale to district scale to make them comparable with the observed yields. For the comparison of spatial patterns, we aggregate the district yields to agro-ecological zones. The humid areas of Tanzania are neglected, because these areas only cover tiny parts of the Tanzanian land surface. Due to the aggregation, the goodness of fit increases retrospectively, because district individual yield anomalies are filtered out (Woodard and Garcia, 2008). All statistical models are applied on district level; we did not aggregate the exogenous variables, because this would lead to information losses due to a reduced variability of the estimation dataset (Garcia et al., 1987). To show the insurance index, we aggregate the yield index from district to regional scale by the arithmetic average (main article Fig. 4).

11.5.1.6 Software

Our statistical models are estimated with the software *R*. We use the package *plm* for the PDMs, the package *lme4* for the RCMs and the package *lmtest* for the statistical tests. The maps are generated with the *R* package *ggplot2*. The process-based model SWIM is written in *fortran*.

11.5.2 Further results and discussion

11.5.2.1 Crop physiological yield assessment by SWIM

The yield level of 1.3 t ha^{-1} is satisfactorily captured by the process-based model SWIM. However, the patterns of low and high yield regions and the inter-annual variability are only poorly (r = 0.05) captured by our process-based model (district-level correlation map in Fig. S.5 and yield levels for each year and district in Fig. S.6). Since process-based models consider only a limited number of processes in their model set-up, they may neglect possibly relevant processes like intercropping, tillage practices. Moreover, an important additional shortcoming of these models might be the lack of management information – e.g., growing season settings, fertilizer application – (Müller et al., 2016). This might be one reason for our unexplained residual yield variability. Another reason might be that the development of these models lags behind the rapidly changing agricultural sector – caused by climate change or technological development – (Rötter et al., 2011). However, Fig. S.5 shows that the water scare regions are covered with higher accuracy than the regions with sufficient water supply. This shows that the model is sensitive for the weather-related influences on crop yields.

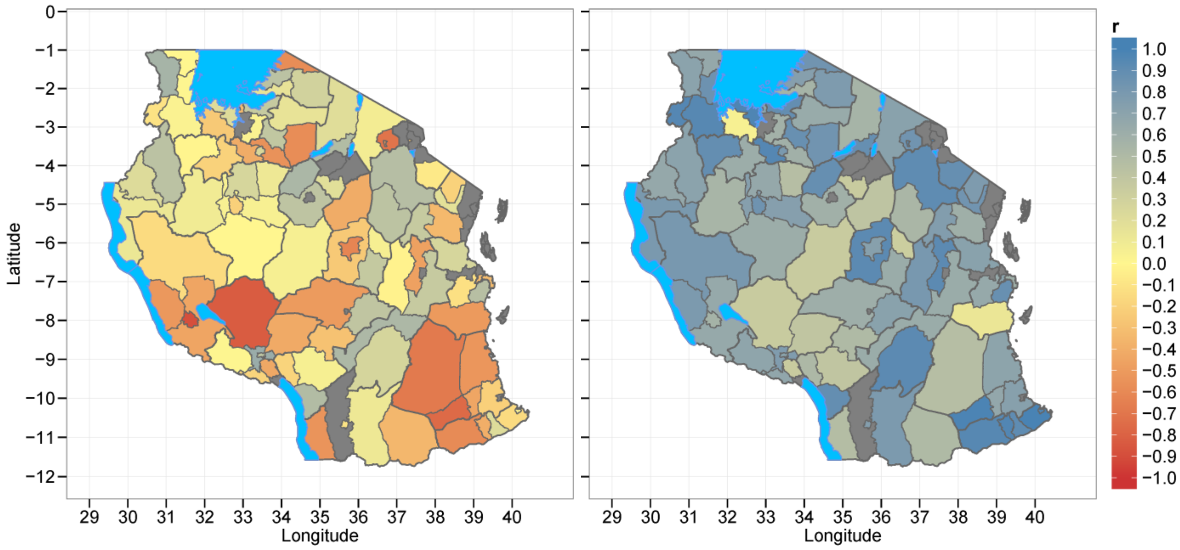


Fig. S.5: Correlation (Pearson's r) of observed maize yields and only process-based modeled yields (left) or process-based and statistical modeled yields (right) at district scale.

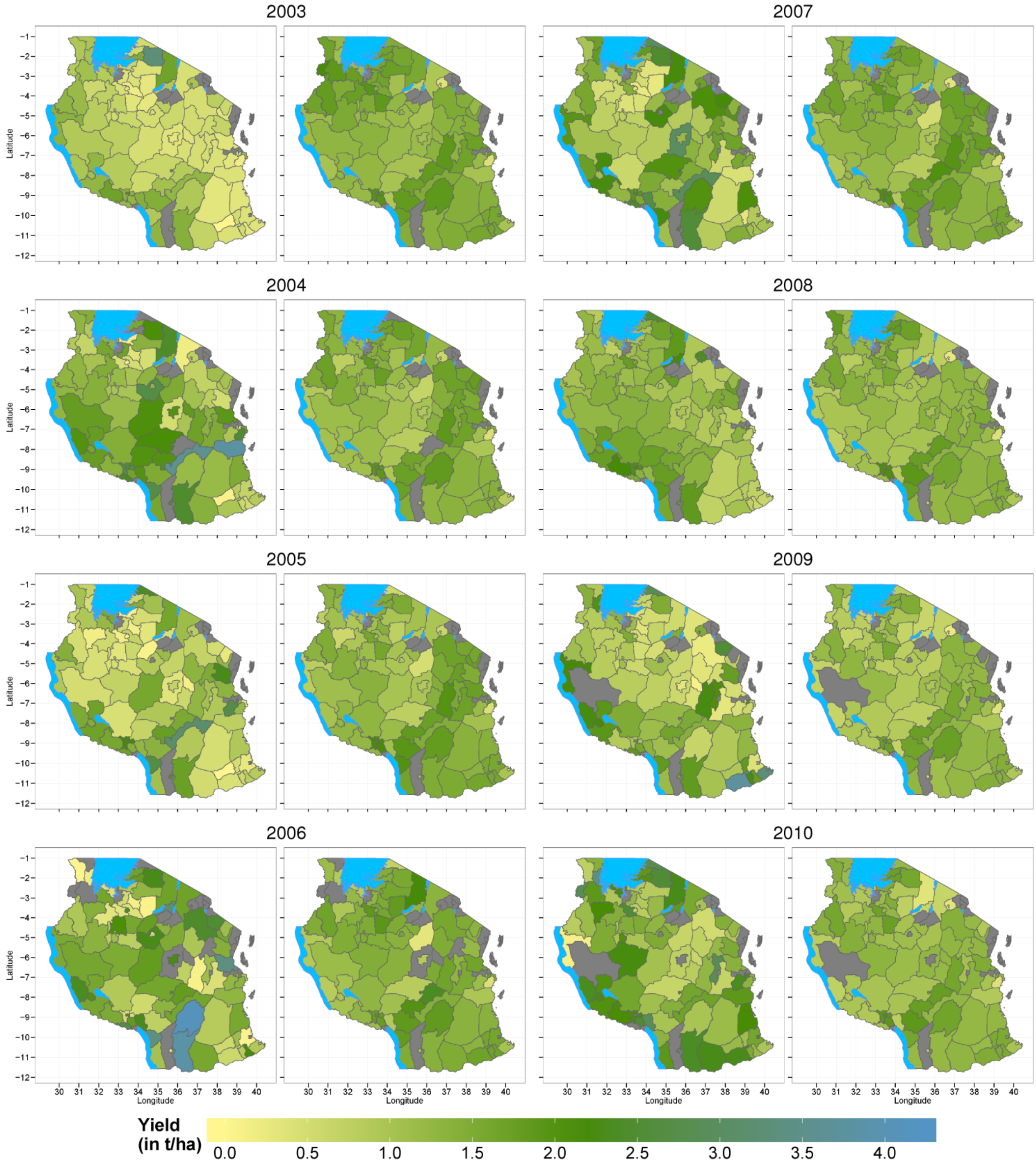


Fig. S.6. Observed (first and third column) and process-based modeled (second and fourth column) yields on district scale for each year of the period 2003-2007.

11.5.2.2 Results for the statistical modeling approach

11.5.2.2.1 Model robustness and selection

The results of the PM^{we}-SM^{nw}-SM^{we2} approach (Eq. S.2) are shown in Tab. S.1 and S.2. For our insurance modeling scheme MAYIS, we use the results of the STSMs row i. The non-weather-related and indirect weather-triggered STSMs (iii) achieve the highest goodness of fit for the estimated yields ($r = 0.92$). The goodness of fit for the STSM with solely non-weather-related impacts (i) is slightly lower ($r = 0.86$), as is also the case for the solely indirect weather-triggered impacts (ii) ($r = 0.78$). The

STSMs show stronger indirect weather-triggered impacts on district scale than the PDM on national scale. This can be explained by the fact that weather-triggered effects (like pest outbreaks or plant diseases) appear rather locally than on a national scale. The generally lower goodness of fit in the validation can be explained by the short (8 years) and sometimes incomplete observed yield time series. Remarkably, the validation results decline when the weather-triggered impacts are also considered. While the STSMs attain a correlation of $r = 0.33$ by including the weather-triggered impacts, by excluding these impacts the correlation rises to $r = 0.38$. This is similar for PDMs and RCMs. Consideration of the solely weather-triggered impacts decreases the validation correlation to $r = 0.04$ for the STSMs, to $r = 0.22$ for the PDMs, and to $r = –0.04$ for RCMs. Thus, we conclude that weather-triggered impacts do not contribute any model robustness. The consideration of both impact factor groups in single PDMs leads also to decreasing estimation and validation power in comparison to the sole consideration of either non-weather-related or weather-triggered factors. The RCMs' goodness of fit is close to that of the STSMs. While the STSMs estimation is higher than for the RCMs, the validation results of the RCMs are slightly higher in comparison to the STSMs. For our insurance calculation, we use the STSMs with only non-weather-related impacts. However, the PDMs with both factor groups are suitable in the case of strong multicollinearity.

Tab. S.1. Correlation of the observed and estimated or validated yields. The columns (STSMs, PDMs, and RCMs) refer to the three statistical methods. All statistical models (SM) are applied on the residuals of the process-based model (PM). The SM considers either only non-weather-related (i) or indirect (second-order) weather-triggered impacts (ii), non-weather-related and indirect weather-triggered impacts in two consecutive models (iii) or in single PDM (iv). The non-weather-driven statistical model is applied without the process-based model (v). The non-weather variables are *maize acreage*, *urea application*, and *paid subsidies on crop production*, while the weather-triggered variables are *SR*, *PREC*, and *VPD*.

	Approach	STSM	PDM	RCM
	Estimation			
i	PM^{we}-SM^{nw} (used for MAYIS)	0.86	0.55	0.80
ii	PM^{we}-SM^{we2}	0.78	0.49	0.68
iii	PM^{we}-SM^{nw}-SM^{we2}	0.92	0.62	0.85
iv	PM^{we}-$SM^{nw\text{-}we2}$	–	0.65	–
	Validation			
i	PM^{we}-SM^{nw} (used for MAYIS)	0.38	0.31	0.43
ii	PM^{we}-SM^{we2}	0.04	0.22	–0.04
iii	PM^{we}-SM^{nw}-SM^{we2}	0.33	0.30	0.40
iv	PM^{we}-$SM^{nw\text{-}we2}$	–	0.17	–

Tab. S.2. R^2 of the observed and estimated or validated yields. The other terms similar to Tab. S.1.

	Approach	STSM	PDM	RCM
	Estimation			
i	PM^{we}-SM^{nw} (used for MAYIS)	0.74	0.30	0.64
ii	PM^{we}-SM^{we2}	0.60	0.24	0.47
iii	PM^{we}-SM^{nw}-SM^{we2}	0.84	0.39	0.72
iv	PM^{we}-$SM^{nw\text{-}we2}$	–	0.42	–
	Validation			
i	PM^{we}-SM^{nw} (used for MAYIS)	0.15	0.10	0.19
ii	PM^{we}-SM^{we2}	0.00	0.05	0.00
iii	PM^{we}-SM^{nw}-$SM^{we2\text{-}log}$	0.11	0.09	0.16
iv	PM^{we}-$SM^{nw\text{-}we2}$	–	0.03	–

In general, our results show that most of the yield variability can be explained by our approach. Sole consideration of the process-based model might initially challenge its usability. However, by consecutively applying the statistical model (Eq. S.2), we are able to explain the remaining yield variability by agronomic management and socio-economy. This justifies the use of the process-based model to calculate weather-attributable impacts. As a result, our combined approach contributes an important tool for the crop modeling community to cover agronomic management and socio-economic impacts and control the non-weather-related yield variability. Our approach allows us to replace district-specific agronomic management information, which is frequently unavailable in particular in SSA (Müller and Robertson, 2014), by a set of regionally available agronomic management and socio-economic variables. Moreover, weather-related yield anomalies coming from other process-based crop models, its ensemble results, or additional observed yield data can be easily incorporated.

11.5.2.2.2 Variable parameters

The parameters of the statistical models driven by the non-weather and indirect weather-triggered variables are shown in Fig. S.7 (PM^{we}-SM^{nw}-SM^{we2}-approch). The non-weather variable *urea supply* has on average a positive yield impact. This means that additional fertilization is positive for the yields. The impact of the *maize acreage* is negative. This is also reasonable since an expansion of the maize acreage is achieved through the cultivation of less suitable land. The impact of *paid agricultural subsidies* is also positive. This is also reasonable, because the increase in improved seeds and fertilizer is positive for the maize yield (Sánchez, 2010). However, the *paid agricultural subsidies* parameter is with 6% added explained yield variability smaller than the other two non-weather-related parameters. This could be interpreted as an indication that these subsidies have a relatively small impact (this is in line with the literature, e.g., Benson et al., 2012)) While the weather variables *VPD* and *SR* have on average a positive yield impact, the *PREC* yield impact is on average close to zero. Moreover, the range of the district individual parameters is higher for the statistical model with indirect weather-triggered variables than for the statistical model with non-weather-related variables. This can be explained by the significantly lower explained yield variability by the statistical model with indirect weather-triggered variables.

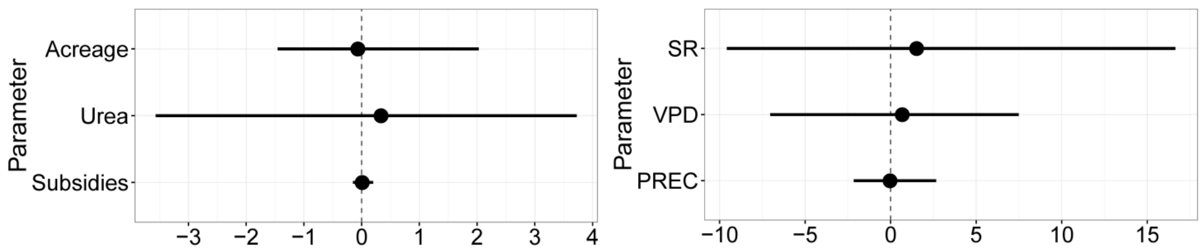


Fig. S.7. Estimated parameters of the non-weather (left) and weather-driven (right) separate time-series model (Eq. S.2). The point is the arithmetic average parameter size of all district models; the lines are the 5% and 95% percentile.

The consideration of a further exogenous variable increases the goodness of fit of the estimation, independently of whether there is an actual contribution by the variable. However, this does not necessarily hold true for the goodness of fit of the validation (Conradt et al., 2016). The variable selection *SR*, *VPD*, *PREC*, TMP_{min} leads to poorer validation results than for the validation with the variables *SR*, *VPD*, *PREC*. This can be explained by the limited degrees of freedom. The variable selection *SR*, *VPD*, *PREC* archieves the best validation goodness of fit. The validation goodness of fit decreases by using the variables selection TMP_{min}, *SR*, *PREC*, and further by the selection TMP_{min}, *VPD*, *PREC*.

Furthermore, we also analyze several other weather and non-weather variables, however, with lower goodness of fit. The additionally tested variables are evapotranspiration (ETP_{TI} by Turc-Ivanov and ETP_H by Haude), *growing degree days* (≥ 8 °C, < 30 °C), *heat degree days* (≥ 30 °C), *temperature normalized solar radiation*, national fertilizer application (*diammonium phosphate* and *calcium ammonium nitrate*), *sprayed area against red locust*, and the Tanzanian *maize price*.

11.5.2.3 Model robustness and uncertainty

11.5.2.3.1 Pre-analysis of significant non-weather-related yield effects

We apply two PDMs on national scale to investigate whether indirect weather-triggered and socio-economic effects are in the residuals (unexplained yield variability) of the process-based model. These residuals are used as the endogenous variable. The exogenous variables of the first PDM are year dummies (to capture year-dependent systemic effects), maize acreage and the weather variables *SR*, ETP_{TI}, and *PREC* (to capture collinear or omitted weather-triggered effects). As result, all year dummy parameters and the acreage are significant with $p \leq 0.01$ and the models provide significant correlation coefficient of $r = 0.40^{***}$ (Pearson correlation; $^{NS}\,p > 0.1$, $^{*}\,p \leq 0.1$, $^{**}\,p \leq 0.05$, $^{***}\,p \leq 0.01$). The consideration of only weather variables, after removing year dummies and the acreage, reduces the r to 0.07^{*} and leaves no significant variables $p \leq 0.01$. Thus, we conclude that only a small effect from weather-triggered impacts remains in the residuals on the national scale. The significant impact of the year dummies indicates an uncontrolled impact of non-weather-related variables.

11.5.2.3.2 Model validity and statistical tests

For statistical models, the estimation method is permissible if the *ordinary least squares assumptions are fulfilled and if no explaining variables are neglected (problem of omitted variable bias). Thus, we conduct* several statistical tests to verify the permissibility of the statistical models, which consider the socio-economic yield impacts. The statistical tests are described by Croissant and Millo (2008) and Baltagi (2005). N*o statistical test exists* for the *problem of omitted variable bias; however,* the regression equation specification error test (RESET) investigates whether quadratic variables are omitted in the model. The RESET shows that quadratic variables are not neglected. Only in 4% of the models, quadratic terms would be beneficial for the model goodness of fit. We also tested several other variable transformations. The chosen log y_t-transformation achieves the highest goodness of fit. The first differences and fixed-effects transformation as well as the untransformed terms achieve lower goodness of fit. The Breusch–Godfrey and the Breusch–Pagan test are applied to test against autocorrelation and heteroscedasticity. In some cases the model residuals are autocorrelated (29%), but mostly they are not (Breusch–Godfrey test). However, autocorrelation can be a problem in macro panels with $T \geq 60$ and $N > T$, but is rather unproblematic in micro panels (Baltagi, 2005*, p. 102-103)*. Since our panel has only a time series length of 8 years and more spatial than the temporal observations, autocorrelation (Breusch–Godfrey test) seems not relevant for our case. Heteroscedasticity appears in 0% (Breusch–Pagan test) of the models. The normal distribution of residuals is tested using the Shapiro–Wilk test. In 4% (Shapiro–Wilk test) of the models, the residuals are not normal distributed. The weather-driven statistical model is not tested, because the results are not further used.

If both weather and management-related factors are estimated in one statistical model, there might be an overlap of these factors due to multi-collinear and/or not clearly assignable processes. These are for instance, temperature and pests & diseases (Rosenzweig et al., 2001) or precipitation and fertilizer efficiency (Alem et al., 2010). In particular, pests & diseases are (at least partly) manageable, but also depend on weather conditions. If such factors are included in a statistical model, this model might not be able to disentangle these collinear processes without a certain uncertainty. Due to the design of our approach, the weather-related part is assessed in the first step (by the process-based model) and only the remaining yield variability is further used for the statistical model assessment. Since process-based models do not face the problem of statistical multicollinearity and its outputs are calculated independently from statistical model outputs, the fraction of overlap should be relatively small in our analysis. For the claim calculation, our approach solely relies on the process-based model, while the statistical model is used to justify the usability of the process-based component. Thus, a limited robustness of the statistical model influences neither the weather-attributable yields nor claim payouts.

11.5.2.3.3 Functional form and variable transformation

The limited degrees of freedom do not allow separating the non-weather and weather-related yield variability within one statistical model. However, we can separately estimate models for both parts. Therefore, we apply the Cobb–Douglas function as a further functional form. The Cobb–Douglas function has been well tested for agronomic and economic applications. To linearize the Cobb–Douglas function, all variables are used as a logarithm. We apply the function to capture either the (first-order) weather (which are assessed by the process-based model in the main approach) or non-weather-related yield impacts. The different variable transformations have been tested previously. The fixed-effects transformation $\left(\log \ddot{y}_t = \log\left(\frac{y_t}{\bar{y}}\right)\right)$ works best, followed by logarithmic transformation ($\log y_t$), followed by first differences $\left(\Delta \log y_t = \log\left(\frac{y_t}{y_{t-1}}\right)\right)$, followed by untransformed values (y_t).

$$\log \ddot{y}_{it} = \log \beta_{0i} + \sum_{j=1}^{J} \beta_{ji} \, \log \ddot{x}_{jit} + \log \ddot{u}_{it}\,, \qquad \text{with } \ddot{y} = \frac{y_{it}}{\bar{y}_i}, \tag{S.4}$$

$\bar{y}$ as arithmetic average of y_t, and respectively for $\ddot{x}$, $\bar{x}$, $\ddot{u}$, and $\bar{u}$.

The utilization of the Cobb–Douglas function with the fixed-effects transformation (Eq. S.4) leads to significantly ($p < 0.01$, Fisher z-transformation) lower goodness of fit in comparison to our used modeling approach (Tab. S.1 row i). The weather-driven statistical model (SM^{we}) attains a correlation of 0.77 (0.10) for the STSMs estimation (validation), 0.64 (0.27) for the PDMs, and 0.72 (0.22) for the RCMs. Fig. S.8 (left) shows that the weather-driven STSM attains a lower goodness of fit than the main (PM^{we}-SM^{nw}) approach. The average deviation from the observed yields (root mean square error) is for the solely weather-driven statistical model 0.53 t ha^{-1} and for the main approach 0.41 t ha^{-1}.

The non-weather-driven statistical model (SM^{nw}) attains a correlation of 0.83 (0.29) for the STSMs, 0.71 (0.50) for the PDMs, and 0.79 (0.49) for the RCMs. Fig. S.8 (right) shows that this modeled yields (STSM, r = 0.83) scatter slightly more around the observed yields than the yields of the main approach (r = 0.86). The root mean square error is 0.46 t ha^{-1} for the solely non-weather-driven statistical model and 0.41 t ha^{-1} for the main approach. The small differences between the non-weather--driven statistical model and the main approach can be explained by the dominate yield impacts of non-weather effects (see also Fig. 3 of the main article: share of weather-related yield losses). The high correlation of non-weather and weather-driven statistical models illustrates a statistical overlap between weather- and non-weather-determined yield variability. This can only be resolved by using a process-based model to capture beforehand the weather-related yield variability, notably under these low degrees of freedom.

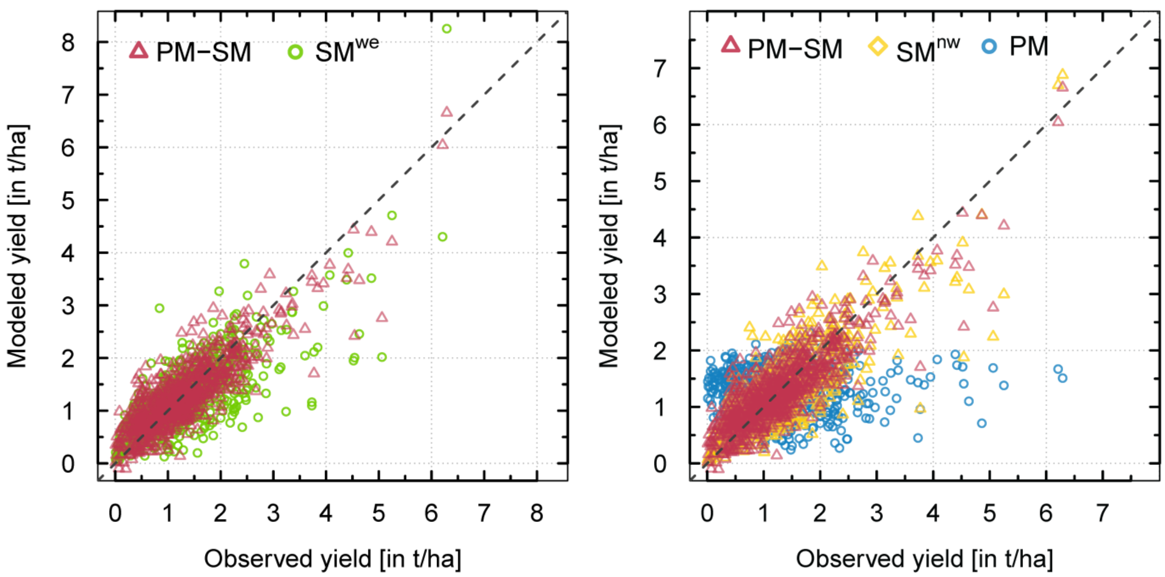


Fig. S.8. Goodness of fit due to the combined application of a process based and a statistical model (PM-SM, red points). Left: the green points show the sole weather-driven statistical model (SM^{we}, Eq. S.4). Right: the yellow points show the sole application of the non-weather-drive statistical model (SM^{nw}, Eq. S.4) and the blue points show the solely applied process-based model (PM).

We also apply our main approach (PM^{we}-SM^{nw}) with the Cobb–Douglas function instead of the logarithmic function as the functional form. The process-based model is used to capture the (first-order) weather-attributable yield impacts and a consecutive statistical model to capture non-weather and indirect weather-triggered influences as in the PM^{we}-SM^{nw}-SM^{we2} approach. The exogenous variables (right side of Eq. S.5) are similarly transformed as in the approach of Eq. S.4. As the endogenous variable, we take the difference of the transformed observed and process-based modeled yields. As transformation, we use the logarithmic fixed-effects. However, this approach achieves the lowest goodness of fit. The correlation of the non-weather-driven statistical models achieve an r of 0.33 (0.18) for the STSMs, of 0.42 (0.33) for the PDMs, and of 0.31 (0.18) for the RCMs (validation results in parentheses). The models for indirect weather-triggered impacts achieve as correlations 0.30 (0.03) for STSMs, 0.34 (0.21) for PDMs, and 0.28 (0.19) for RCMs. Thus, we conclude that the Cobb–Douglas function is less suitable for the cropping conditions in Tanzania.

$$\log \ddot{y}_{it} - \log \ddot{y}_{it}^{PM} = \log \beta_{0i} + \sum_{j=1}^{J} \beta_{ji} \log \ddot{x}_{jit} + \log \ddot{u}_{it} \qquad \text{(S.5)}$$

11.5.3 Implementation of the insurance scheme

11.5.3.1 Implementation scheme and innovative products

A successful and sustainable insurance implementation scheme should include as main components i) a framework and structural aspects, ii) operational aspects and iii) should be innovative (Herbold, 2014). (i) The framework and structural aspects contain a network of different stakeholders. The major three stakeholder groups are the farmers, the insurance companies and the government. Moreover, intermediaries and financial institutions as well as a loss assessment entity should be integrated in the network. Within the network, the stakeholders have different roles: the government will be responsible for the legal framework. Since crop yield insurances will be unaffordable for (smallholder) farmers, a co-financing from the government or international funds like the Green Climate Fund will be helpful to make the premiums affordable. The insurance company (actual risk taker) and the government should bear the risk together in a public–private partnership. In particular in developing countries, the lack of risk transfer tools and risk takers often prevent the establishment of an insurance solution. The risk taker should spread the risk by indemnifying other perils across more countries due to a joint market penetration. The (local) risk taker could also transfer the systemic part of the risk to re-insurance companies, because it is likely that a local insurance company cannot bear the entire systemic risk (of e.g., severe droughts across an entire country). The loss assessment service, which would potentially use the suggested modeling scheme, has to be independent from the risk taker and political interest to avoid tampering or conflicts of interests (similar to the claim adjuster of classical indemnity insurances). It should bundle the technical expertise and maintain the necessary database. The intermediary and financial institutions should have access to the individual smallholder farmers and should facilitate a cost-efficient distribution of the insurances. The intermediary and financial institutions are, for instance, local banks, extension officers or input retailer, but also new mobile money transfer services are thinkable. However, it is also possible that the local insurer will have direct link to the individual farmers. In Fig. S.9, we have visualized the main elements of the structural framework for crop insurance implementation.

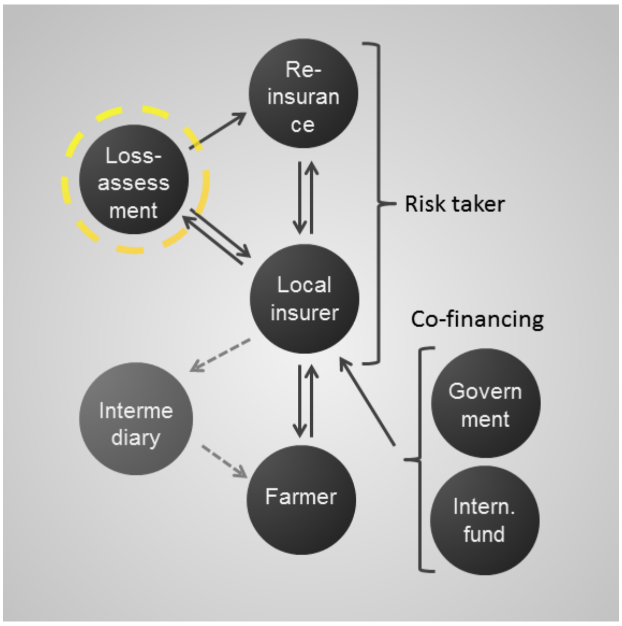


Fig. S.9: Structural framework for crop insurance implementation in developing countries. The loss assessment service (dashed yellow circle) could potentially be our modeling scheme.

ii) The second requirement addresses the operational aspects. Here, it is important that once the structural framework is put in place, loss determination is carried out consistently. For the loss determination, it is essential to have local and modeling expertise to conduct the loss assessments. This could be a challenge in SSA where crop modeling is slowly and sparsely represented within higher education resulting in little qualified staff. Thus, the engagement with local universities and recruitment of trained experts could be helpful. As a co-benefit this would also facilitate capacity building for data surveys, statistical analysis and crop modeling. For the operation, the premium calculation will be conducted before planting, while the actual loss assessment and the claim determination will be conducted after the harvest or a yield failure. The premium policies are calculated on basis of the average yield loss within a historical period (Leblois and Quirion, 2013) and adjusted whenever new (historical) data are available. For the claim calculation, our approach should assess yield losses and calculate the indemnities immediately after the harvest by the process-based model.

Finally, iii) the innovativeness is important for the implementation success of the insurance scheme (Herbold, 2014). In particular in connecting smallholder farmers with the insurer and for the loss assessment without claim adjusters, innovative solutions are needed. For connecting farmers, making use of mobile money transfer services to ensure direct claims payouts without temporal delay can strengthen the acceptance of smallholder farmers in insurances. For the loss assessment, the literature provides several potential loss determination triggers for weather index insurances, which are discussed in the next section.

11.5.3.2 Weather index insurance solutions

Weather index insurances are designed to indemnify crop yields of smallholder farmer groups within similar agro-ecological conditions. Within these groups, farmers achieve the same claims per unit land in case of yield loss independently from the actual yield loss on their field (Sarris, 2013). The claims are determined by indexes, because of (otherwise too) high transaction costs for the claim determination carried out by claim adjusters as it is usually done for indemnity-based insurances (Meze-Hausken et al., 2009). The trigger for claim determination and payoffs is often based on a cumulated precipitation index or a related parameter. The advantage of such indexes is that they are easy to understand (e.g., in case of no rainfall, insurance claims will be disbursed). This is helpful for the implementation of the insurance and the acceptance by smallholder farmers in SSA (Patt et al., 2010). However, the relationship between weather indexes and actual crop yields is often not linear (Lobell et al., 2011; Rowhani et al., 2011). Thus, precipitation indexes are only partly suitable as trigger for yield losses assessments (Herbold, 2014). Besides the basic precipitation indexes, more sophisticated indexes, like water stress or drought indexes, are also applied. Moreover, there are also approaches, which utilize more parameters in addition to precipitation (e.g., GDD, frost indexes) within statistical crop models (Okhrin et al., 2013; Xu et al., 2010) and have a clear focus to capture the extreme downside yield anomalies (Conradt et al., 2015). Finally, satellite imagery data or process-based models are also applied in some cases as index for yield anomalies (Leblois and Quirion, 2013). Regardless of whether a yield or a precipitation index will be applied, the yield loss index should be highly correlated with the actual yield losses. The basis risk is defined as the risk that index measurements will not match actual farm-individual losses. In general, there are three main types of basis risk (Norton et al., 2012): the spatial basis risk results from the bias to capture the local conditions (e.g., distance to the next weather station). The temporal basis risk originates from a temporal mismatch between the considered time period and actual growing period at the farm level (Dalhaus and Finger, 2016). The design basis risk originates from the insurance contract design and thus from the loss model itself (Elabed et al., 2013). At district level, our approach achieves a good coverage of the actual yield variability. This indicates a low basis risk at this level. Nevertheless, we are aware that the basis risk will be higher at the farm level.

11.5.3.3 Comparison between a precipitation index and MAYIS

To address the design basis risk, we conduct a comparison of our modeled yield index and a precipitation index as it is frequently used for weather index insurances. Fig. S.10 shows the correlation of our suggested approach with the observed yields on district level (right). In comparison, the application of a precipitation index leads to a much lower coverage of the observed yields (left). This corroborates the hypotheses that the relationship between yield and precipitation is too complex to be captured by a simple index using a single weather variable. For the comparison in Fig. S.10, we use the correlation coefficient (r), because the correlation coefficient also gives information about negative correlations in comparison to the R^2. The modeled MAYIS yield index has an overall performance of r = 0.86 and no negative correlation values. In contrast, the correlation of the cumulated precipitation index (aggregated over the growing season) explains the observed yield variability only by r = 0.10. The range of the individual districts is in the latter case is between r = 0.94 and r = –0.80 (95, 75, 25, and 5% percentile: r = 0.79, 0.34, –0.25, –0.54). This shows that in most cases precipitation is not a good proxy for the actual yield variability in Tanzania. In case of high negative correlations, the precipitation index would lead to a claim payoff in case of above-average yields and no payoffs in case of severe yield losses. Such a mechanism would immediately erode the trust of the farmers in crop insurances and reduce the farmers' willingness to pay of farmers (Hill et al., 2013; Patt et al., 2009). However, the correlation of precipitation and yield might be higher in more water-scarce regions (in other countries) and thus, such precipitation indexes are relevant as trigger of yield loss determination (Berg et al., 2009; Carter et al., 2016).

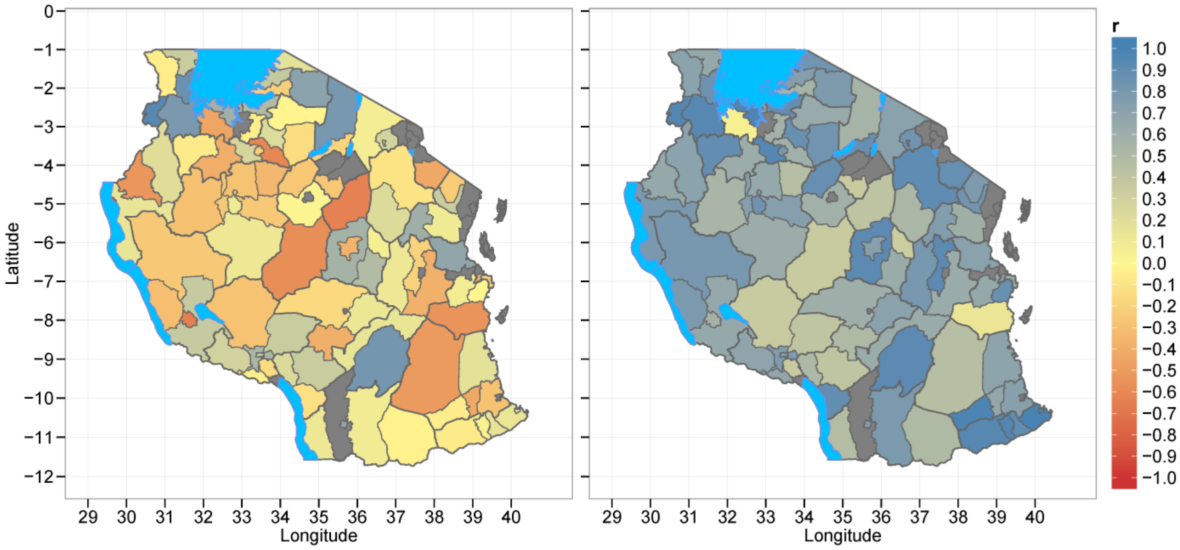


Fig. S.10: Correlation (Pearson's r) of observed maize yields and a precipitation index (left) or modeled yields from the MAYIS approach (right) at district scale. The precipitation index is the cumulated precipitation over growing season, which are typically used as trigger for weather index insurances.

11.5.4 References

Alem, Y., Bezabih, M., Kassie, M., Zikhali, P., 2010. Does fertilizer use respond to rainfall variability? Panel data evidence from Ethiopia. Agric. Econ. 41, 165–175.
Asseng, S. et al. , 2013. Uncertainty in simulating wheat yields under climate change. Nat. Clim. Chang. 3, 827–2.
Baltagi, B.H., 2005. Econometric Analysis of Panel Data, Third edit. ed. John Wiley & Sons Ltd.
Benson, T., Kirama, S.L., Selejio, O., 2012. The supply of inorganic fertilizers to smallholder farmers in Tanzania evidence for fertilizer policy development. IFPRI Discuss. Pap. 1230, 1–48.
Berg, A., Quirion, P., Sultan, B., 2009. Weather-Index Drought Insurance in Burkina-Faso: Assessment of Its Potential Interest to Farmers. Weather. Clim. Soc. 1, 71–84.
Carter, M.R., Cheng, L., Sarris, A., 2016. Where and how index insurance can boost the adoption of improved agricultural technologies. J. Dev. Econ. 118, 59–71.
Cassman, K.G., Dobermann, A., Walters, D.T., Yang, H., 2003. Meeting Cereal Demand While Protecting Natural Resources and Improving Environmental Quality. Annu. Rev. Environ. Resour. 28, 315–358.
Castellvi, F., Perez, P.J., Stockle, C.O., Ibanez, M., 1997. Methods for estimating vapor pressure deficit at a regional scale depending on data availability. Agric. For. Meteorol. 87, 243–252.
Chipanshi, A., Zhang, Y., Kouadio, L., Newlands, N., Davidson, A., Hill, H., Warren, R., Qian, B., Daneshfar, B., Bedard, F., Reichert, G., 2015. Evaluation of the Integrated Canadian Crop Yield Forecaster (ICCYF) model for in-season prediction of crop yield across the Canadian agricultural landscape. Agric. For. Meteorol. 206, 137–150.
Conradt, S., Finger, R., Bokusheva, R., 2015. Tailored to the extremes: Quantile regression for index-based insurance contract design. Agric. Econ. 46, 537–547.
Conradt, T., Gornott, C., Wechsung, F., 2016. Extending and improving regionalized winter wheat and silage maize yield regression models for Germany: Enhancing the predictive skill by panel definition through cluster analysis. Agric. For. Meteorol. 216, 68–81.
Croissant, Y., Millo, G., 2008. Panel data econometrics in R: The plm package. J. Stat. Softw. 27.
Dalhaus, T., Finger, R., 2016. Can Gridded Precipitation Data and Phenological Observations Reduce Basis Risk of Weather Index–Based Insurance? Weather. Clim. Soc. 8, 409–419.
Deininger, K., Savastano, S., Xia, F., 2017. Smallholders' land access in Sub-Saharan Africa: A new landscape? Food Policy 67, 78–92.
Dewitte, O., Jones, A., Spaargaren, O., Breuning-Madsen, H., Brossard, M., Dampha, A., Deckers, J., Gallali, T., Hallett, S., Jones, R., Kilasara, M., Le Roux, P., Michéli, E., Montanarella, L., Thiombiano, L., Van Ranst, E., Yemefack, M., Zougmore, R., 2013. Harmonisation of the soil map of Africa at the continental scale. Geoderma 211–212, 138–153.
Dillon, B., Barrett, C.B., 2017. Agricultural factor markets in Sub-Saharan Africa: An updated view with formal tests for market failure. Food Policy 67, 64–77.
Elabed, G., Bellemare, M.F., Carter, M.R., Guirkinger, C., 2013. Managing basis risk with multiscale index insurance. Agric. Econ. 44, 419–431.
FAO, 2015. Crop water information: Maize. URL http://www.fao.org/nr/water/cropinfo_maize.html (accessed 10.17.16).
FAO Crop Calendar, 2010. United Republic of Tanzania – Maize. URL http://www.fao.org/agriculture/seed/cropcalendar/ (accessed 10.17.16).
FAO Stat, 2013. FAO Stat database collections: Maize prices for Tanzania (2003-2010). URL http://faostat3.fao.org/ (accessed 10.17.16).
Folberth, C., Gaiser, T., Abbaspour, K.C., Schulin, R., Yang, H., 2012. Regionalization of a large-scale crop growth model for sub-Saharan Africa: Model setup, evaluation, and estimation of maize yields. Agric. Ecosyst. Environ. 151, 21–33.
Folberth, C., Skalský, R., Moltchanova, E., Balkovič, J., Azevedo, L.B., Obersteiner, M., van der Velde, M., 2016. Uncertainty in soil data can outweigh climate impact signals in global crop yield simulations. Nat. Commun. 7, 11872.
Folberth, C., Yang, H., Gaiser, T., Abbaspour, K.C., Schulin, R., 2013. Modeling maize yield responses to improvement in nutrient, water and cultivar inputs in sub-Saharan Africa. Agric. Syst. 119, 22–34.
Gaiser, T., de Barros, I., Sereke, F., Lange, F.-M., 2010. Validation and reliability of the EPIC model to simulate maize production in small-holder farming systems in tropical sub-humid West Africa and semi-arid Brazil. Agric. Ecosyst. Environ. 135, 318–327.
Garcia, P., Offutt, S.E., Pinar, M., Changnon, S.A., 1987. Corn yield behavior: Effects of technological advance and weather conditions. J. Clim. Appl. Meteorol. 26, 1092–1102.
Gornott, C., Wechsung, F., 2016. Statistical regression models for assessing climate impacts on crop yields: A validation study for winter wheat and silage maize in Germany. Agric. For. Meteorol. 217, 89–100.
Harmsen, K., 2000. A modified mitscherlich equation for rainfed crop production in semi-arid areas: 1. Theory. NJAS - Wageningen J. Life Sci. 48, 237–250.

Herbold, J., 2014. New Approaches to Agricultural Insurance in Developing Economies, in: Köhn, D. (Ed.), Finance for Food: Towards New Agricultural and Rural Finance. Springer Berlin Heidelberg, pp. 199–217.
Hill, R.V., Hoddinott, J., Kumar, N., 2013. Adoption of weather-index insurance: learning from willingness to pay among a panel of households in rural Ethiopia. Agric. Econ. 44, 385–398.
IFPRI, 2015. Agro-ecological zones of Sub-Saharan Africa (8-class). URL https://harvestchoice.org/data/aez8_clas (accessed 10.17.16).
Iizumi, T., Ramankutty, N., 2015. How do weather and climate influence cropping area and intensity? Glob. Food Sec. 4, 46–50.
ILRI, 2005. Soil type distribution map of Tanzania. URL http://data.ilri.org/geoportal/ %3E Tanzania Soil (accessed 10.17.16).
Jayne, T.S., Mather, D., Mason, N., Ricker-Gilbert, J., 2013. How do fertilizer subsidy programs affect total fertilizer use in sub-Saharan Africa? Crowding out, diversion, and benefit/cost assessments. Agric. Econ. 44, 687–703.
Kaufmann, R.K., Snell, S.E., 1997. Biophysical Model of Corn Yield: Integrating Climatic and Social Determinants. Am. J. Agric. Econ. 79, 178–190.
Krysanova, V., Hattermann, F., Huang, S., Hesse, C., Vetter, T., Liersch, S., Koch, H., Kundzewicz, Z.W., 2015. Modelling climate and land-use change impacts with SWIM: lessons learnt from multiple applications. Hydrol. Sci. J. 60, 606–635.
Krysanova, V., Wechsung, F., Arnold, J., Srinivasan, R., Williams, J., 2000. Soil and Water Integrated Model: User manual. Pik Rep. 69, 1–243.
Leblois, A., Quirion, P., 2013. Agricultural insurances based on meteorological indices: realizations, methods and research challenges. Meteorol. Appl. 20, 1–9.
Lobell, D.B., Bänziger, M., Magorokosho, C., Vivek, B., 2011. Nonlinear heat effects on African maize as evidenced by historical yield trials. Nat. Clim. Chang. 1, 42–45.
Lukeba, J.L., Vumilia, R.K., Nkongolo, K.C.K., Mwabila, M.L., Tsumbu, M., 2013. Growth and leaf area index simulation in maize (Zea mays L.) under small-scale farm conditions in a Sub-Saharan African region. Am. J. Plant Sci. 4, 575–583.
MAFSC, 2010. Agricultural statistics. Ministry of Agriculture, Food Security and Cooperatives. URL http://www.kilimo.go.tz/agricultural statistics/ (accessed 10.17.16).
McClung, C.R., 2014. Making hunger yield. Science 344, 699–700.
Meze-Hausken, E., Patt, A., Fritz, S., 2009. Reducing climate risk for micro-insurance providers in Africa: A case study of Ethiopia. Glob. Environ. Chang. 19, 66–73.
Minot, N., Benson, T., 2009. Fertilizer subsidies in Africa - are vouchers the answer? IFPRI Discuss. Pap.
Müller, C., Elliott, J., Chryssanthacopoulos, J., Arneth, A., Balkovic, J., Ciais, P., Deryng, D., Folberth, C., Glotter, M., Hoek, S., Iizumi, T., Izaurralde, R.C., Jones, C., Khabarov, N., Lawrence, P., Liu, W., Olin, S., Pugh, T.A.M., Ray, D., Reddy, A., Rosenzweig, C., Ruane, A.C., Sakurai, G., Schmid, E., Skalsky, R., Song, C.X., Wang, X., de Wit, A., Yang, H., 2016. Global Gridded Crop Model evaluation: benchmarking, skills, deficiencies and implications. Geosci. Model Dev. Discuss. 1–39.
Müller, C., Robertson, R.D., 2014. Projecting future crop productivity for global economic modeling. Agric. Econ. 45, 37–50.
NBS, 2012. National sample census of agriculture – Small holder agriculture. Volume II: crop sector – National report. Minist. Agric. Food Secur. Coop. Minist. Livest. Dev. Fish. Minist. Water Irrig. Minist. Agric. Livest. Environ. Zanzibar, Prime Minist. Off. Reg. Adm. .
Nkonya, E., Mwangi, W., 2004. The economic rationale of recycling hybrid seeds in Northern Tanzania. East. Afr. J. Rural Dev. 20, 113–124.
Norton, M.T., Turvey, C., Osgood, D., 2012. Quantifying spatial basis risk for weather index insurance. J. Risk Financ. 14, 20–34.
Okhrin, O., Odening, M., Xu, W., 2013. Systemic Weather Risk and Crop Insurance: The Case of China. J. Risk Insur. 80, 351–372.
Patt, A., Peterson, N., Carter, M., Velez, M., Hess, U., Suarez, P., 2009. Making index insurance attractive to farmers. Mitig. Adapt. Strateg. Glob. Chang. 14, 737–753.
Patt, A., Suarez, P., Hess, U., 2010. How do small-holder farmers understand insurance, and how much do they want it? Evidence from Africa. Glob. Environ. Chang. 20, 153–161.
Pauw, K., Thurlow, J., 2011. Agricultural growth, poverty, and nutrition in Tanzania. Food Policy 36, 795–804.
Rosenzweig, C., Elliott, J., Deryng, D., Ruane, A.C., Müller, C., Arneth, A., Boote, K.J., Folberth, C., Glotter, M., Khabarov, N., Neumann, K., Piontek, F., Pugh, T. a M., Schmid, E., Stehfest, E., Yang, H., Jones, J.W., 2014. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. Proc. Natl. Acad. Sci. 111, 3268–3273.
Rosenzweig, C., Iglesias, A., Yang, X.B., Epstein, P.R., Chivian, E., 2001. Climate change and extreme weather events: Implications for food production, plant diseases, and pests. Glob. Chang. Hum. Heal. 2, 90–104.
Rötter, R., Van de Geijn, S.C., 1999. Climate change effects on plant growth, crop yield and livestock. Clim. Chang. 43, 651–681.

Rötter, R.P., Carter, T.R., Olesen, J.E., Porter, J.R., 2011. Crop–climate models need an overhaul. Nat. Clim. Chang. 1, 175–177.

Rowhani, P., Lobell, D.B., Linderman, M., Ramankutty, N., 2011. Climate variability and crop production in Tanzania. Agric. For. Meteorol. 151, 449–460.

Sánchez, P.A., 2010. Tripling crop yields in tropical Africa. Nat. Geosci. 3, 299–300.

Sarris, A., 2013. Weather index insurance for agricultural development: introduction and overview. Agric. Econ. 44, 381–384.

Schlenker, W., Lobell, D.B., 2010. Robust negative impacts of climate change on African agriculture. Environ. Res. Lett. 5, 1–8.

Tanzania Meteorological Agency, 2007. Daily weather data 1970–2006

Thornton, P.K., Jones, P.G., Alagarswamy, G., Andresen, J., 2009. Spatial variation of crop yield response to climate change in East Africa. Glob. Environ. Chang. 19, 54–65.

Tittonell, P., Giller, K.E., 2013. When yield gaps are poverty traps: The paradigm of ecological intensification in African smallholder agriculture. F. Crop. Res. 143, 76–90.

van der Velde, M., Folberth, C., Balkovič, J., Ciais, P., Fritz, S., Janssens, I. a., Obersteiner, M., See, L., Skalský, R., Xiong, W., Peñuelas, J., 2014. African crop yield reductions due to increasingly unbalanced nitrogen and phosphorus consumption. Glob. Chang. Biol. 20, 1278–1288.

Vitousek, P.M., Naylor, R., Crews, T., David, M.B., Drinkwater, L.E., Holland, E., Johnes, P.J., Katzenberger, J., Martinelli, L.A., Matson, P.A., Nziguheba, G., Ojima, D., Palm, C.A., Robertson, G.P., Sanchez, P.A., Townsend, A.R., Zhang, F.S., 2009. Nutrient Imbalances in Agricultural Development. Science (80). 324, 1519–1520.

Ward, P.S., Florax, R.J.G.M., Flores-Lagunes, A., 2014. Climate change and agricultural productivity in Sub-Saharan Africa: A spatial sample selection model. Eur. Rev. Agric. Econ. 41, 199–226.

Weedon, G.P., Balsamo, G., Bellouin, N., Gomes, S., Best, M.J., Viterbo, P., 2014. Data methodology applied to ERA-Interim reanalysis data. Water Resour. Res. 50, 7505–7514.

Westengen, O.T., Ring, K.H., Berg, P.R., Brysting, A.K., 2014. Modern maize varieties going local in the semi-arid zone in Tanzania. BMC Evol. Biol. 14, 1.

Woodard, J.D., Garcia, P., 2008. Weather derivatives, spatial aggregation, and systemic risk: Implications for reinsurance hedging. J. Agric. Resour. Econ. 33, 34–51.

World Bank, 2016. Tanzania national panel survey. URL http://go.worldbank.org/EJMAC1YDY0 (accessed 10.17.16).

Xu, W., Filler, G., Odening, M., Okhrin, O., 2010. On the systemic nature of weather risk. Agric. Financ. Rev. 70, 267–284.

You, L., Ringler, C., Wood-Sichra, U., Robertson, R., Wood, S., Zhu, T., Nelson, G., Guo, Z., Sun, Y., 2011. What is the irrigation potential for Africa? A combined biophysical and socioeconomic approach. Food Policy 36, 770–782.

You, L., Rosegrant, M.W., Wood, S., Sun, D., 2009. Impact of growing season temperature on wheat productivity in China. Agric. For. Meteorol. 149, 1009–1014.

Selbständigkeitserklärung

Hiermit erkläre ich, die Dissertation selbstständig und nur unter Verwendung der angegebenen Hilfen und Hilfsmittel angefertigt zu haben. Ich habe mich anderwärts nicht um einen Doktorgrad beworben und besitze keinen entsprechenden Doktorgrad. Ich erkläre, dass ich die Dissertation oder Teile davon nicht bereits bei einer anderen wissenschaftlichen Einrichtung eingereicht habe und dass sie dort weder angenommen noch abgelehnt wurde. Ich erkläre die Kenntnisnahme der dem Verfahren zugrunde liegenden Promotionsordnung der Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin. Weiterhin erkläre ich, dass keine Zusammenarbeit mit gewerblichen Promotionsbearbeiterinnen/Promotionsberatern stattgefunden hat und dass die Grundsätze der Humboldt-Universität zu Berlin zur Sicherung guter wissenschaftlicher Praxis eingehalten wurden.

Declaration:

I hereby declare that I completed the doctoral thesis independently based on the stated resources and aids. I have not applied for a doctoral degree elsewhere and do not have a corresponding doctoral degree. I have not submitted the doctoral thesis, or parts of it, to another academic institution and the thesis has not been accepted or rejected. I declare that I have acknowledged the doctoral degree regulations which underlie the procedure of the Faculty of Life Sciences of Humboldt-Universität zu Berlin. Furthermore, I declare that no collaboration with commercial doctoral degree supervisors took place and that the principles of Humboldt-Universität zu Berlin for ensuring good academic practice were abided by.

Berlin, den ______________ ____________________________

Datum Unterschrift